LISTEN TO THE WORD MEANING OR THE VOCAL TONE: SENSITIVITY TO EMOTIONAL CUES AMONG EUROPEAN AMERICAN AND CHINESE CHILDREN

A Dissertation

Presented to the Faculty of the Graduate School

of Cornell University

In Partial Fulfillment of the Requirements for the Degree of

Doctor of Philosophy

by

Yang Yang

August 2017

LISTEN TO THE WORD MEANING OR THE VOCAL TONE: SENSITIVITY TO EMOTIONAL CUES AMONG EUROPEAN AMERICAN AND CHINESE CHILDREN

Yang Yang, Ph. D.

Cornell University 2017

People from varied cultural backgrounds differ in their attention to particular aspects of emotional cues. Whereas semantic content explicitly expresses feelings, vocal tone conveys implicit information regarding emotions. This dissertation examined the attention to different emotional cues in European American, Chinese urban, and Chinese rural children. Participants were 121 European American, 120 Chinese urban, and 130 Chinese rural children (4-9 years old). Half were girls and half were boys in each cultural group. They played two computer games in which they listened to spoken words and judged the pleasantness of the word meaning while ignoring the vocal tone (Meaning game) or judged the pleasantness of the vocal tone while ignoring the word meaning (Tone game). Results showed that European American children paid more attention to word meanings than did Chinese children. In contrast, Chinese children, especially 8-9-year-olds, attended more to vocal tones than did their European American counterparts. The two groups of Chinese children performed similarly in the tasks in general. The results are discussed in terms of cultural similarities and differences in children's emotion understanding, as well as for their practical implications in inter-cultural emotional communication.

BIOGRAPHICAL SKETCH

Yang Yang was born in Beijing, China and received her Bachelor's degrees in Medicine and Psychology at Peking University. Subsequently, she attended Harvard Graduate School of Education to study the major of mind, brain, and education, and obtained her Master's degree in Education there. Her research experiences at Harvard and Cornell cultivated her research interests in questions including how culture influences parents' socialization process, how parents' behavior and parent-child interaction affect children's emotion understanding, and the role of children's emotion knowledge on their later socioemotional development.

ACKNOWLEDGEMENTS

I would not be able to finish my dissertation without the guidance of my committee members, help from lab members, CV Starr Fellowship from Cornell EAP program, and support from family and friends.

I would like to express my greatest gratitude to my advisor, Dr. Qi Wang. Her advices in both research and professional development guided my way towards a good researcher, and her resources in the lab allowed me to explore the research questions that I'm interested in. I am very grateful to Dr. Anthony Ong for advising me to be "selective" when I was a first-year graduate student. Dr. Jim Booth is extremely approachable and helpful when I have questions about statistics. His course cultivated my interests in the mathematical mechanisms under statistical methods and led me to pursue a minor in statistics. Although Dr. Felix Thoemmes is not on my committee, his encouragement is critical for my confidence to survive in graduate school. I also appreciated Dr. Tamar Kushnir to share her resources in the Sciencenter where I recruited most my American participants for my dissertation. I would also express my gratitude to Dr. Li Wang for her help on my data collection in China.

My lab mates, Jessie Koh and Qingfang Song provided valuable guidance and feedback on my projects. My research assistants in the Culture & Social Cognition Laboratory have contributed to my dissertation project considerably by working on the experiment materials and recruiting participants. I would also like to thank the friends in my department, Youjeong Park and Carissa Kang for sharing experiences on writing and

defending dissertation, and Alice Zhao for her company and help during my data collection.

I want to thank my family who serve as the secure base for me to pursue what I want and always encourage me with their best wishes. They also helped tremendously on my data collection in China. And finally, I want to thank my husband, Yue Yu, for his intellectual advice, technical support, proof reading on my dissertation, and for always being there cheering me up.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

Figure Page

CHAPTER 1

INTRODUCTION

In the process of globalization, communication among people from different cultures is unavoidable. In the past decades, there are an increasing number of Chinese immigrants in the US. In many situations, European Americans and Chinese have to communicate with each other in their life. The awareness of cultural difference in emotion processing can facilitate inter-cultural communication. This dissertation aimed to explore the cultural similarities and differences in emotion processing, and provide implication on emotion communication among people from different cultural backgrounds.

An emotional utterance contains multiple aspects of information, such as verbal content of the emotional words as well as the vocal tone. People from various cultures differ in their attention to particular aspects. For example, European Americans pay more attention to verbal content, whereas East Asians are more sensitive to contextual information such as vocal tones (Ishii, Reyes, & Kitayama, 2003). In each culture, adults usually transmit cultural values to children through emotion socialization (Wang, 2013). As a result, children may learn a culturally specific pattern of attention to emotional cues at an early age. This project investigates children's sensitivity to emotional cues in different cultural contexts.

In the following sections of introduction, I first introduced different nonverbal channels of emotion communication. Second, I reviewed the Hall's theory about low- and high-context cultures and summarized empirical studies that demonstrated Asian cultures are high-context whereas American culture is low-context. Differences in cognition and emotion communication between low-context and high-context cultures were also reviewed. Third, I proposed that adults might transmit their culture-unique pattern of behaviors to children through socialization process and summarized studies on emotion socialization. Finally, cross-cultural studies on children's emotion processing were reviewed.

Nonverbal emotion communication

Human can communicate emotions clearly through language, but a great amount of emotional information is conveyed through multiple channels of nonverbal cues. The most significant nonverbal cue when communicating emotions is facial expression, which conveys rich information about emotions (Awasthi & Mandal, 2015).

Facial expression

Facial expressions are generated by different combinations of facial muscles, which are controlled by neural network (Ekman, 1972). Because of this biological nature, some elements of facial expressions are universal among people who are in literate or preliterate cultures, blind or with eyesight, and in Western or Eastern countries (Ekman et al., 1987; Izard, 1994; Matsumoto & Willingham, 2009). On the other hand, facial

expressions are also influenced by socio-cultural environment from the beginning of individuals' life (Awasthi & Mandal, 2015). People judge facial expressions more accurately if the emotion is expressed by members from their own culture than by those from other cultures (Elfenbein & Ambady, 2002). Cultures have different display rules about facial expressions of emotions, which constrain when individuals can express emotions, which emotions are appropriate to express in a given situation, who can express such emotions to whom, and how intensely the emotions should be expressed (Ekman, 2007).

Gestures and distance

Beyond facial expression, many other human's behaviors carry emotional cues as well, but some of those are more implicit than facial expression. For example, gestures during speech can also help people understand a speaker's emotions. Gestures can provide information for people to differentiate between positive and negative emotions, and between high arousal and low arousal emotions (Castellano, Villalba, & Camurri, 2007). Specifically, basic gestural form features, including handedness, hand shape, and motion direction, are related to emotions (Kipp & Martin, 2009). Additionally, distance also plays a role in emotion communication (Buck & Miller, 2014). Moderate personal distances are best for communication of basic emotions, including happiness, sadness, fear, anger, surprise, and disgust. Several emotions, such as love, anger, and sympathy are communicated well at intimate distances. Social and moral emotions such as pride,

guilt, and shame are usually communicated at longer social and public distance. When communicating emotions like trust, gratitude, and respect, interpersonal synchrony and mutually contingent responsiveness are usually involved (Buck & Miller, 2014). When emotions are judged based on multiple cues, such as face-gesture-speech, the recognition rate is higher than the rate in situations when emotion is judged based on any one type of cues (e.g., only face) (Kessous, Castellano, & Caridakis, 2010).

Vocal tone

Vocal tone is a particularly critical message in emotion communication, because it exists in all emotional speech and people can receive emotional cues from vocal tones even when a speaker is not in sight. Vocal tone is also evolutionarily important due to its function in individuals' survival (Scherer, 2003). When encountering predators, vocal tone is ideal to pass information over long distance. It is adaptive to use vocal tone to warn family or friends (in fear) immediately without moving towards them closely, or to threaten the enemy (in anger) in a safe distance. Multiple elements in vocal tones are related to emotions. Listeners' judgments of emotions are largely based on mean fundamental frequency (F0), standard deviation of F0, mean energy, duration of voiced period, proportion of spectral energy up to 1000Hz, and spectral drop-off (Banse & Scherer, 1996). Among these factors, the range of fundamental frequency influences the judgments the most. Specifically, narrow range of fundamental frequency is interpreted as sadness, whereas wide range of fundamental frequency is correlated with high arousal

emotions. Loudness in its own does not affect judgments of emotions when the voice qualities are the same (Yanushevskaya, Gobl, & Ní Chasaide, 2013). The intensity and speed are also indicative of emotions. Speech with high speech intensity is usually judged as negative emotions. Fast speech is usually inferenced as joy, whereas slow speech often leads to judgment of sadness (Scherer, 2003). The time needed to recognize different emotions according to speech prosody is also various in different emotions. Fear is usually recognized fastest, but happiness is recognized relatively slowly (Pell & Kotz, 2011; Rigoulot, Wassiliwizky, & Pell, 2013). Recognizing emotion from auditory cues involves multiple brain structures distributed between both left and right hemisphere in processing specific auditory features related to certain emotions. Among the involved brain structures, the right inferior frontal regions is considered the most critical area, which works together with more posterior regions in the right hemisphere, left frontal regions, and subcortical structures (Adolphs, 2002). However, research shows that the judgment of emotions based on vocal tones is generally less accurate than judgment based on facial expression (Ekman, 1994; Scherer, 1999, 2003; Scherer, Banse, & Wallbott, 2001). This result suggests that vocal tone may be less salient and more implicit than facial cues, and may be harder to detect than facial expressions during emotion communication.

Emotion communication in high-context vs. low-context cultures

Researchers across different disciplines, such as psychology and anthropology, have suggested that in many Western cultures, verbal content plays the primary role of conveying information during communication, whereas in East Asian cultures more information is conveyed by contextual cues (Ambady, Koo, Lee, & Rosenthal, 1996; Hall, 1976; Markus & Kitayama, 1991). Hall (1976) proposed the concepts of *low-context* and *high-context* cultures. In low-context cultures, such as European American culture, individuals' thoughts need to be explicitly expressed in words in order to be understood. Therefore, verbal content contains the major proportion of information in communication in these cultures and a smaller proportion of information is carried in contextual cues. People in low-context cultures need to communicate in more explicit details because they share less common information and are less able to infer what is not said. In contrast, in high-context cultures, such as East Asian cultures, people are deeply involved with each other and share a great amount of information in common. The shared information is not necessary to be explicitly said during communication. Therefore, individuals' ideas can be understood without explicit expression in words as long as enough contextual information is given. Thus, in such cultures, contextual cues play a larger role in communication, whereas explicit verbal content conveys a relatively smaller proportion of information, compared to low-context cultures. Consequently, the communication in low-context cultures is more direct, less implicit, and more

informative, whereas the communication in high-context cultures is generally more indirect, more implicit, and highly depends on contexts.

Reasons for high-context and low-context cultures

One possible reason for East Asian cultures (e.g., China, Korea, Japan) to be high-context is the particular characters and sentence structures used in Chinese written language (Hall, 1976). Hall (1976) noticed that Chinese written language is thirty-five hundred years old and has not changed very much in the past three thousand years. This shared written language unifies people in China, Korea, and Japan. Context is needed in many activities in these cultures. For example, when looking up words in a dictionary, the radicals are necessary to find a word. Hall (1976) gave an example that in order to find the word, star (星), in a dictionary one must know it's under the radical of sun (日).

Empirical studies also support that East Asia is more high-context than American culture. A study conducted in China, Korea, and the US showed that Chinese and Korean people scored higher than Americans on the aspects that were consistent with Hall's conceptualization of high-context cultures, such as higher involvement with other people and more implicit communication (Kim, Pan, & Park, 1998). In a study on politeness in Korea and the US, Koreans' choice of politeness strategies was more affected by relational cues (i.e., contextual information), whereas Americans' choice of politeness strategies was influenced by the verbal content (Ambady, et al., 1996). In linguistics field, Kashima and Kashima (1998) examined whether a pronoun can be dropped as a

subject of a sentence across 39 languages in 71 cultures. They found that in English, pronouns are usually necessary to be included in a grammatically correct sentence. These structural characteristics of language require individuals in English-speaking cultures to explicitly express their ideas in verbal communication, which is a low-context form of communication. In contrast, in languages in East Asian cultures, such as Chinese, Japanese, and Korean, pronouns are generally optional in a sentence structure. Sentences in these languages are often ambiguous; and communication in these cultures largely depends on contextual information.

Consequences of high-context and low-context cultures

One consequence of East Asian cultures being high-context is that communication in these countries becomes more indirect. Speakers in theses cultures only provide part of information and expect the listeners to fill the remaining pieces of information (Hall, 1976; Salleh, 2005). For example, a friend says she is hungry but does not have time to buy food before finishing her work. Listeners in low-context culture may think the friend will keep hungry until she completes her work. In a high-context culture, the listener may get the implied information that her friend is trying to ask whether she can buy food for her. It feels demanding to ask a favor directly in high-context cultures, whereas indirect communication may allow individuals to achieve their needs when the listeners can pick up the implicit cues. The feature of indirectness in high-context communication is supported by empirical studies as well. In a study investigating perceptions of subtle

racist speech, both Asian Americans and European Americans rated the perceived harm of direct and indirect racist speech uttered by a white person to a non-white person (Leets, 2003). Results showed that Asian Americans rated the indirect racist speech as the most harmful, whereas European Americans evaluated the harmfulness mainly based on the verbal content of the message (Buck & Miller, 2014). These results suggest that Asian Americans are more sensitive to indirect cues than European Americans.

The differences between high-context and low-context cultures even shape different styles of attention and cognition between Easterners and Westerners (Nisbett, Peng, Choi, & Norenzayan, 2001). Studies employing a variety of cognitive tasks consistently demonstrated that individuals in Japan, a high-context culture, are more sensitive to contextual information than people in the United States, a low-context culture. For example, in one study, American and Japanese participants were shown animated video clips of underwater scenes containing a focal fish and background objects (Masuda & Nisbett, 2001). When asked to describe what they saw in the scenes, Japanese participants referred more to the background objects and less to the focal fish than Americans. Additionally, Japanese participants were more likely to recognize the focal fish that they had seen before if the fish was presented against the original background than in situations when the focal fish was shown in a different background. However, Americans' performances on recognizing the focal fish were the same regardless of whether the background was original or novel. These findings suggest that Japanese were

more likely to attend to background information and their performances in tasks were influenced by the background to a greater extent than Americans. This cultural difference in attention to contextual information has also been found in tasks which are more abstract (Kitayama, Duffy, Kawamura, & Larsen, 2003) or more social (Masuda et al., 2008). Masuda and his colleagues (2008) found that Japanese participants', but not Westerners', judgments of the central person's emotion were influenced by the surrounding people's emotions. Japanese participants also looked at the surrounding people more than Westerners. Not only in Japan, but also in China, studies have shown that East Asians are more sensitive to contextual information than Americans. Eye-tracking studies revealed that Chinese participants fixated more on the background, while North American participants fixated more on the focal object when they were viewing pictures of naturalistic scenes (Chua, Boland, & Nisbett, 2005; Lu, Daneman, & Reingold, 2008). In sum, East Asians generally pay more attention to contextual cues than do Americans.

A third impact of high- and low-context culture is the importance of nonverbal information in communicating emotions. Silence does not mean absence of thoughts; instead it may convey rich information (Ramsey, 1998). It may indicate politeness to a senior by waiting for the senior's speech first; it may show respect to group harmony; it may also mean disagreement with previous speech (Dsilva & Whyte, 1998; Ramsey, 1998). People in high-context cultures tend to express their opinions through multiple

nonverbal channels discussed previously in the emotion communication section, such as gestures, facial expressions, and vocal tones. For example, people may show their disagreements by slightly shaking heads, clearing throats, or looking at the person who did inappropriate behaviors for a long time. Vocal tone is an important nonverbal channel of emotion communication which conveys contextual information. As expected, East Asians are more sensitive to vocal tones of emotional utterances than Americans. Kitayama and his colleagues conducted a series of studies on adults' attention to the meanings of emotional words and the vocal tones of utterances in multiple cultures with Stroop tasks (Ishii, et al., 2003; Kitayama & Ishii, 2002). In these studies, participants listened to emotional words with either a congruent (i.e., a pleasant word with a pleasant tone; an unpleasant word with an unpleasant tone) or an incongruent vocal tone (i.e., a pleasant word with an unpleasant tone; an unpleasant word with a pleasant tone). They were asked to judge the pleasantness of the utterance either by the word's meaning while ignoring the vocal tone, or by the vocal tone while ignoring the word meaning. Results showed that Japanese participants were more interfered by the vocal tone than by the verbal content, whereas for Americans the interference was greater when they were asked to ignore word's meaning than the condition in which they were asked to ignore the vocal tone. The results consistently suggest that East Asians automatically pay more attention to the context (vocal tone) of an emotional utterance and less attention to the verbal content of the emotional words, compared to European Americans. The same cultural

difference in the patterns of attention to emotional cues has been shown in American and Asian participants with a wide range of ages (18 – 78 years old). However, whether children have acquired the sensitivity to emotional cues from vocal tones has yet to be investigated.

Emotion socialization and socialization regarding contextual information

Adults in various cultures transmit their cultural values about emotions and culturally unique pattern of behaviors regarding emotions to children through socialization process. Emotion socialization aims to facilitate children's emotion competence (Denham, Bassett, & Wyatt, 2007; Eisenberg, Cumberland, & Spinrad, 1998). In studying emotion socialization, it is necessary to consider the role of culture on parental socialization goals and how children's emotion competence is defined. Parents' culturally-shaped values and goals guide their socialization behaviors (Cole & Tan, 2007). Thus, there are cultural differences in emotion socialization.

Emotion socialization in different cultures

In several Western countries, such as the United States and Germany, parents emphasize children's autonomy and independence. Expression of self is encouraged in these cultures. Emotion expression is considered as assertion of self and is valued in these cultures. Individuals need to express themselves explicitly in order to achieve their goals (Markus & Kitayama, 2001). Parents in these cultures model various emotions by expressing their own emotions appropriately. Through parents' modeling, children learn

important knowledge about situations which elicit certain emotions and also skills to regulate emotions. Parents' expressions of a particular group of emotions promote children's general emotion competence (Halberstadt, Fox, & Jones, 1993; Valiente et al., 2004), whereas parents' lack of expression of emotions impedes children's emotional expressiveness and emotion knowledge in general (Suveg, Zeman, Flannery-Schroeder, & Cassano, 2005). In addition, parents in these cultures also teach, share, discuss emotions with their children (i.e., emotion coaching). Talking about feelings and explaining causes and consequences of emotions assists children in expressing emotions, understanding emotions, and managing their emotions (Denham, Zoller, & Couchoud, 1994; Garner, Jones, Gaddy, & Rennie, 1997). Discussion about emotional experiences between parents and children helps children to learn great knowledge about emotional expressions, situations, and causes (Denham, Renwick-DeBardi, & Hewes, 1994; Dunn, Brown, & Beardsall, 1991; Dunn, Slomkowski, Donelan, & Herrera, 1995). Finally, parents' reaction to children's emotions is another important avenue to socialize emotion. When children express negative emotions, European American parents usually accept children's negative emotions, encourage them to express and experience these emotions, and provide emotional comfort and assistance. However, reactions as ignoring, minimizing, or punishing children's negative emotions are viewed harmful to children's emotion competence (Gottman, Katz, & Hooven, 1997).

In contrast, in East Asian cultures, people value group harmony, hierarchical relationships, and interdependence (Matsumoto, 1991). Emotion expression is considered disruptive to group harmony, so is usually discouraged in these cultures (Wang, 2003). With these cultural values, Chinese parents talk less about causes and consequences of emotions with their children compared to European American parents. Instead, Chinese parents teach emotion display rules, promote sensitivity to others' emotions, and emphasize proper behavioral conduct (Chan, Bowes, & Wyver, 2009; Wang, 2001; Wang & Fivush, 2005; Wang, Leichtman, & Davies, 2000). In response to children's negative emotions, parents in these cultures have less strong preference on different types of reactions than European American parents. For example, Chinese mothers have similar preference between minimizing children's negative emotions and encouraging children's emotion expressions (Chan, et al., 2009; Tao, Zhou, & Wang, 2010).

Given that the socialization practices are deeply shaped by culture, we expect that the cross-cultural differences on values of contextual information, such as the importance of vocal tone in communicating emotions, should be vertically transmitted from one generation to the next through cultural-specific socialization practices. In line with this view, since European American culture is low-context, European American parents aim to help their children to articulate and explicitly express their emotions in order to achieve what they need (Chao, 1995). On the other hand, Chinese parents often expect their children to attend to others' emotions and to infer others' feelings without being

explicitly told in order to align to high-context communication style (Chao, 1995; Chen, 2000; Wang, Hutt, Kulkofsky, McDermott, & Wei, 2006). Chinese parents are also less likely to explicitly explain the emotions than European American parents (Doan & Wang, 2010). As a result, compared to European American children, Chinese or Chinese immigrant children have a lower level of explicit emotion knowledge, such as the ability to explain the situations that elicit specific emotions (Doan & Wang, 2010; Wang, 2008). Nevertheless, Chinese parents more often use subtle ways to simply point out other's emotions without engaging in extensive discussion about feelings. In these subtle ways, Chinese parents help children to read others' "face color" (脸色, a Chinese idiom for the facial expression of emotions) and to behave accordingly (Wang, 2013). Consequently, Chinese children may learn to be more sensitive to implicit cues of emotions and to be better able to infer others' emotions merely from contextual cues (e.g., vocal tone) without being told. However, Chinese children's implicit emotion knowledge, namely, the ability to infer others' emotions from implicit cues, has yet to be studied.

Socialization of attention in low-context and high-context cultures

Cross-cultural differences also exist in socialization practices regarding attention towards contextual cues for non-emotional information. One study conducted among Japanese children suggested that parents transmitted their culturally specific pattern of attention to children, although children only show that pattern when their parents complete the task together with them (Senzaki, 2013). In this study, Japanese and

Canadian children aged 4-9 years watched the videos of underwater scenes described previously. When asked to recall what they saw in the videos, older children (aged 7-9) talked more in general than younger children, but did not show any cultural difference in their references to either the focal fish or the background. However, in a follow-up study, when children discussed and recalled the videos together with their mothers, the older Japanese children referred to more information about the background but less information about the focal fish than did older Canadian children. No cultural difference was found in the recall among the younger children. These studies suggested that children have not gained the culture-specific pattern of attention at the age of 4 – 9 years when they were asked to independently recall the videos. Nevertheless, the cultural difference in the pattern of attention has emerged for children aged 7-9 when they were scaffold by their parents. Children younger than 6 years did not show a culturally specific pattern of attention may be due to the specific features of the task used in these studies. Firstly, the task in Senzaki's studies (2013) required cognitive skills other than attention, such as memory and verbal descriptions, which were difficult for young children. The challenges in these additional skills may veil the different patterns of attention among preschoolers. Combined with what we discussed above, parents from different cultures have distinct socialization goals regarding children's attention to contextual information for both emotional and non-emotional information. However, it is still unclear whether children

learn cultural-specific patterns of attention to contextual cues and use them when they are alone.

Cross-cultural studies on children's emotion processing

Emotion experience and expression in different cultures

Consistent with cultural values and emotion socialization shaped by culture, children's emotion expression and experience vary across cultures. Chinese children were reported to experience fewer emotions than American children (Porter et al., 2005). In observational studies, Japanese children express anger less frequently than American children across different contexts of assessment (Zahn-Waxler, Friedman, Cole, Mizuta, & Hiruma, 1996). Japanese infants showed fewer emotional responses to a routine inoculation than European American infants (Lewis, Ramsay, & Kawakami, 1993). This cultural difference in emotion expressivity was more robust between Chinese and European American children than the difference between Japanese and European American children. Camras (1998) found that Chinese infants were less emotionally expressive than European American and Japanese children (Camras, et al., 1998). In another study, Korean preschoolers took part in several emotion-eliciting tasks; their positive and negative emotion expressions were observed. Results showed that Korean children expressed sadness or exuberance less frequently than European American children when participating in the emotion-eliciting tasks (Louie, Oh, & Lau, 2013). In

general, East Asian children experience or express emotions less frequently than European American children.

Emotion knowledge in different cultures

Cross-cultural studies on emotion knowledge have also shown differences between Asian and European American preschoolers' emotion knowledge, as mentioned above in emotion socialization section. In one study, Chinese and European American preschoolers were presented with 20 short stories in which the protagonist experienced discrete emotions. They were asked to judge which emotion the protagonist experienced in each story. Chinese preschoolers performed poorer on this emotion situation knowledge task than European American preschoolers (Wang, 2003; Wang, et al., 2006). In another task assessing emotion knowledge, children were asked to describe situations that could elicit certain discrete emotions. Chinese and Chinese American children also scored lower on this emotion knowledge task than European American children across preschool years (Doan & Wang, 2010; Wang, 2008; Wang, et al., 2006). This cultural difference in children's emotion knowledge is at least partially due to the difference in parental emotion socialization practices in Chinese and American cultures (Doan & Wang, 2010). Culture shapes the style of mother-child conversation, which in turn influences children's emotion situation knowledge.

Emotion recognition in different cultures

However, studies do not show that Asian children perform poorer on emotion recognition tasks. Markham and Wang (1996) examined Chinese and Australian children's emotion recognition. In this study, children saw facial expressions of 6 basic emotions (i.e., happiness, sadness, fear, anger, disgust, and surprise) and were asked to label them or give example situations. Results revealed that Chinese children performed better than Australian children on recognition of all emotions except for happy. Markham and Wang (1996) argued that Chinese children's higher accuracy of emotion recognition may be due to culturally shaped emotion socialization from Chinese parents. Chinese culture emphasizes group harmony. The emphasis on group harmony may lead to the individuals in this culture being more sensitive to other group members' implicit emotional expressions. As a result, members in this culture may learn to recognize emotions even from attenuated displays of emotions. Studies have shown that children's sensitivity to emotion expressions with various intensity continues to improve between 5 and 10 years of age (Gao & Maurer, 2010), but no cross-cultural studies have been done yet to investigate the cultural difference in children's sensitivity to emotional expressions with various intensity. In another study, Canadian children were asked to judge basic emotions expressed on Caucasian and Asian faces. Results showed that Canadian children categorized expressions of fear and surprise better from Asian faces, whereas

categorized expressions of disgust better from Caucasian faces (Gosselin & Larocque, 2000).

In terms of recognizing emotions from auditory cues, one study conducted in Canada and Japan showed that school-aged children from both cultures could accurately recognize emotions based on auditory cues in songs (Adachi, Trehub, & Abe, 2004). In this study, both children and adults from Canada or Japan listened to songs sung by Canadian school-aged children. Participants were asked to rate how happy or sad each song sounded. Results showed that both children and adults from either Canada or Japan recognized the intended emotions in the songs. Regardless of culture, children showed higher accuracy than adults given that the songs were sung by children who were at the same age as theirs. Canadian children rated happy songs more accurately than sad songs, whereas Japanese children rated sad songs more accurately than happy songs. However, this study did not directly compare Japanese and Canadian children's accuracy on recognizing emotions from auditory cues. Further, in this study, Japanese children listened to foreign songs in English, so it might not reveal their potential ability to detect emotional cues from auditory information. Although this study demonstrated that children can detect emotional cues from auditory information, studies with direct comparison between East Asian and Western Children are needed.

The importance of preschool years and middle childhood

Preschool years is a critical period when children' knowledge of emotion develops rapidly. They have learned knowledge about emotional expressions and situations that elicit emotions (Denham & Couchoud, 1990). They also extend this knowledge to understand other people's emotions. They start being involved in discussions about causes and consequences of emotions and also learning some complex aspects of emotion, such as display rules (Denham, 1986; Gross & Harris, 1988). Usually by the end of preschool years, children begin to take into account personalized information, such as personal characteristics and the unique situation the person is in, when understanding others' emotion (Gnepp, 1989). Studies have shown that preschoolers, even as young as 3 years old, have learned culture-specific pattern of emotion related behaviors (Friedlmeier, Corapci, & Cole, 2011; Trommsdorff & Friedlmeier, 2009). On the other hand, in early middle childhood, children start school and their interaction with peers increase dramatically. They have more opportunities to learn and practice their skills regarding recognizing and regulation emotions during the interactions with peers. Furthermore, their more advanced cognitive skills enable them to develop conceptual knowledge about more complex emotions (Lagattuta & Thompson, 2007; Thompson, 1989). Thus, middle childhood is a significant time when emotion knowledge is broadened in scope (Saarni, Campos, Camras, & Witherington, 2006).

Therefore, preschool years and middle childhood are great periods to study children's emotions, and the role of culture in children's emotional development.

Present study

Considering that Chinese and American cultures differ in the emphasis on contextual information, parents in these two cultures have different culturally shaped socialization goals and practices, Chinese and American children may develop different sensitivity to relatively implicit emotional cues. Since Chinese parents put more emphases on group harmony and socialize children to attend to implicit cues of others' emotions, Chinese children may have greater sensitivity to implicit emotional cues than their American counterparts.

It is worth noticing that China has carried out a large-scale social and economic reform since 1978, and Chinese people have been introduced and exposed to massive Western values and ideologies since then. During these past four decades, China's economy was reformed toward a market economy which resulted in increases in income and its variation, as well as in international trade and foreign investment. In addition to economic changes, the reform also brought social and cultural interactions. For example, in 1979, there were 1953 international students in China, whereas the number increased to 397,635 in 2015 (Ministry of Education of the People's Republic of China, 2015). Thanks to the number of foreigners living in China, Chinese people have dramatically increased opportunities to interact with Western people. Furthermore, the modern

technology also allows Chinese to absorb Western cultural values through internet, social media, movies, music, and so on. Therefore, Chinese people's values are to some extent westernized especially for young people in large cities in urban areas.

However, there are significant urban-rural differences in economic and social development in China. The economic and social reform has been majorly happening in urban cities. People in rural China have much more limited exposure to Western people or values, compared to their urban counterparts. Due to the imbalanced social and economic development in China in the past decades, a significant cultural discrepancy in people's beliefs and behaviors has been built between rural and urban areas. The higher levels of absorption of Western culture in cities have substantially influenced parenting practices and child behavioral patterns compared with the more traditional lifestyle in the countryside. Research has shown that rural parents in China were more likely to maintain parenting attitudes that were consistent with the traditional values, compared to urban parents in China (Chen, Bian, Xin, Wang, & Silbereisen, 2010). It is possible that Chinese children in rural areas are more likely to be socialized with traditional Chinese values than children in urban China.

This study examined children's sensitivity to emotional cues using the Stroop task developed by Kitayama and his colleagues (Ishii, et al., 2003; Kitayama & Ishii, 2002), and we focused on three groups of children: European American, Chinese urban, and Chinese rural children. Given that people in rural China usually hold traditional Chinese

values, I hypothesized that Chinese rural children would be more sensitive to the vocal tone and less sensitive to the word meaning of the emotional utterance than European American children. Specifically, in the Stroop task, Chinese rural children were predicted to be more interfered by the vocal tone, whereas European American children would have a greater interference effect by the verbal content. In terms of children from urban China, it is possible that they are influenced by Western cultures considerably so that their performance would be similar to European American children's. However, it is also possible that the cultural values regarding emotions are too ingrained to be changed by the exposure of Western values, and Chinese urban children, like Chinese rural children, would be more sensitive to vocal tones and less sensitive to word meanings than European American children.

CHAPTER 2

METHOD

Design

This study employed a cross-sectional mixed design. We tested children from three cultural groups, European American, Chinese urban, and Chinese rural areas. All children were tested for the same two tasks, Meaning task and Tone task. The order of tasks was counterbalanced across children.

Participants

A total of 371 European American (EA), Chinese urban (CU), and Chinese rural (CR) children aged 4-9 years participated in the study. Children in the three samples were matched in terms of age and gender. The specific numbers of boys and girls in each age group and each cultural group are presented in Table 1. EA Children were recruited in local schools and a science center in upstate New York, among children whose parents were both European Americans. Children in urban China were recruited in schools in the urban areas (within Road San Huan) in Beijing, whereas children in rural China were recruited from schools in a village near the border between Beijing and Hebei Province. Parents of Chinese participants were both Chinese. Children in the schools came from the residential areas in which the school is located.

Table 1. The numbers of participants by gender, age, and culture.

	Age 4-5		Age 6-7		Age 8-9	
	Girls	Boys	Girls	Boys	Girls	Boys
EA	20	21	20	20	20	20
CU	20	20	19	20	21	20
CR	20	22	21	22	25	20

Procedure

We modified the Stroop task developed by Kitayama and his colleagues (Isshi et al., 2003) to test children's sensitivity to emotional content and emotional tones in spoken words. In a quiet room, children were introduced to the tasks as two "computer games" in which they were asked to listen to emotional utterances and respond to them. The utterances were spoken in the participants' native language. Children were told that they would hear many words that were different in both their emotional meanings and the emotional tones of voice. Some of the spoken words were congruent utterances (i.e., pleasant word meaning with pleasant tone, unpleasant word meaning with unpleasant tone), while other of the auditory stimuli were incongruent utterances (i.e., pleasant word meaning with unpleasant vocal tone, unpleasant word meaning with pleasant vocal tone). In the Meaning game, children were instructed to judge the meaning of each word as good or bad while ignoring the sound of the utterance (to-be-ignored aspect of the

utterance in the Meaning task). In the Tone game, children were instructed to judge the utterance as either "sound happy" or "sound sad or mad" merely according to the vocal tone while ignoring the meaning of the word (to-be-ignored aspect of the utterance in the Tone task). To make sure that children understood the concept of tone, the experimenter asked children before the Tone game, "Do you know what tone means?" and gave one example of the participant's name with a happy tone and a sad tone to explain the concept of tone. The order of the two games was counterbalanced.

To facilitate children's motor response, we used large JellyBean buttons (2.5-in. diameter) instead of computer keys in this task. On each button, there was either a pleasant cartoon face or an unpleasant cartoon face. Children were asked to use one finger for each button. They needed to press the button with a pleasant face when they judged the utterance as pleasant, otherwise to press the button with an unpleasant face. They were asked to play as quickly as they could while also maintaining response accuracy. Response time was measured from the offset of the stimuli in milliseconds.

All participants were presented with 44 trials (12 practice trials, followed by 32 experiment trials) in each game, and 88 trials in total. In the first four practice trials in each game, the to-be-ignored aspect of the utterances was neutral. For example, at the beginning of the Meaning game, we presented four utterances with pleasant or unpleasant meaning and neutral tone to facilitate participants' understanding of the game. In contrast, the first four practice trials in the Tone game were utterances with neutral

meaning and pleasant or unpleasant tone. The later 8 practice trials contained the same format of utterances in the testing trials - utterances with pleasant or unpleasant tone and pleasant or unpleasant meaning. The second set of 4 practice trials contained utterances with congruent word meanings and vocal tones. The last 4 practice trials presented utterances with incongruent word meanings and vocal tones. The 32 testing trials included utterances with congruent or incongruent word meaning and vocal tone. The order of the utterances in the practice trials was fixed, while the order of the utterances in the testing trials was randomized.

Before each utterance, a "+" appeared in the center of the screen on each trial to warn participants that the utterance was coming. The emotional utterances were only presented from the headphone rather than the screen. After each utterance, a yellow smiley face and a red frowny face were presented on the screen to remind participants to press one of these two buttons. After a button was pressed, participants were given both visual and auditory feedback during the practice trials. If the answer was correct, a green check mark, " √ " appeared on the screen, along with a correct answer sound, "Di", presented in the headphone. Otherwise, a red question mark, "?", and a wrong answer sound, "Zzzz", were presented. During the testing trials, participants did not receive feedback anymore.

Materials

Following Ishii, Reyes and Kitayama (2003), we developed stimulus utterances in four steps. First, we selected 28 pairs of translation-equivalent Mandarin and English words from McAuthur Communication Development Inventory (MCDI) which contains a list of words commonly produced by toddlers. The selected words include nouns, adjectives, and verbs with different emotional meanings, which were 14 pairs of words with pleasant meaning and 14 pairs with unpleasant meaning (Table 2). The 28 Chinese words were rated by 21 Chinese native speakers and the 28 English words were rated by 10 English native speakers in terms of the pleasantness of word meaning (*1= very unpleasant, 7 = very pleasant*).

Table 2. Selected words with pleasant or unpleasant meanings for recording.

Recorded words			
Pleasant meaning		Unpleasant meaning	
English	Mandarin	English	Mandarin
New	新的	Mad	生气
Pretty	漂亮	Scared	害怕
Happy	高兴	Hate	讨厌
Like	喜欢	Sick	生病
Clean	干净	Noisy	吵
Flower	花儿	Hurt	痛
Cute	可爱	Tired	累了
Good	好	Cold	冷
Help	帮忙	Sad	伤心
Cookie	饼干	Dark	天黑
Lollipop	棒棒糖	Dirty	脏
Smile	笑	Bad	不好
fun	有趣的	Broken	坏了
Yummy	好吃	Hungry	饿了

Second, we trained one female Mandarin-English bilingual research assistant to read all the English and Mandarin words in two different tones of voice, including a pleasant tone in which the speakers read it with a smooth and round tone, and an unpleasant tone in which a harsh and constricted tone was used. It yielded 112 utterances in total (28 words x 2 languages x 2 tones).

Third, the 112 utterances resulting from the second step were listened to and judged by 11 bilingual adults in terms of the pleasantness of the vocal tones (*1= very unpleasant, 7 = very pleasant*). Finally, we selected the utterances to be used in the experiment based on the ratings of the pleasantness of both vocal tones (judged in the

third step) and word meanings (judged in the first step). The final set of stimuli includes 64 utterances in testing trials 8 utterances x 2 language (Mandarin and English) x 2 meanings (pleasant and unpleasant) x 2 vocal tones (pleasant and unpleasant), and 16 utterances (2 utterances x 2 language x 2 meanings x 2 vocal tones) in the practice trials. In addition, the bilingual research assistant also recorded four neutral utterances for each task, including 4 utterances with pleasant or unpleasant meaning but neutral tone for the Meaning game, and another four utterances with pleasant or unpleasant tone but neutral meaning for the Tone game. The final set of words are presented in Table 3, Table 4, and Table 5. As can be seen in Table 6, in the final set of utterances used in testing trials, vocal tone was manipulated independently of language and word meaning, and vocal tone and word meaning were equally extreme in the two languages.

Table 3. Selected utterances in the practice trials of Meaning game.

Practice Trials for Meaning Game			
Pleasant meaning		Unpleasant meaning	
English	Mandarin	English	Mandarin
Neutral tone			
Yummy	好吃	Cold	冷
Cookie	饼干	Dark	天黑
Pleasant tone			
New	新的	Bad	不好
Help	帮忙	Hurt	痛
Unpleasant tone			
Happy	高兴	Sick	生病
Clean	干净	Noisy	吵

Table 4. Selected utterances in the practice trials of Tone game.

Practice Trials for Tone Game			
Pleasant tone		Unpleasant tone	
English	Mandarin	English	Mandarin
Neutral words			
Table	桌子	Couch	沙发
Chair	椅子	Window	窗子
Pleasant words			
New	新的	Happy	高兴
Help	帮忙	Clean	干净
Unpleasant words			
Bad	不好	Sick	生病
Hurt	痛	Noisy	吵

Table 5. Selected utterances in the testing trials.

Testing Trials			
Pleasant meaning		Unpleasant meaning	
English	Mandarin	English	Mandarin
Pretty	漂亮	Mad	生气
Like	喜欢	Scared	害怕
Flower	花儿	Hate	讨厌
Cute	可爱	Tired	累了
Good	好	Sad	伤心
Lollipop	棒棒糖	Dark	天黑
Smile	笑	Broken	坏了
fun	有趣的	Hungry	饿了

Note. All words were presented with both pleasant and unpleasant tones.

Table 6. Mean pleasantness ratings for the vocal tones and word meanings of Chinese and English utterances used in the testing trials

	Word meaning			
	Pleasant		Unpleasant	
Language	Pleasant vocal tone	Unpleasant vocal tone	Pleasant vocal tone	Unpleasant vocal tone
	Vocal tone ratings[a]			
Chinese	5.445(.430)	2.990(.330)	5.364(.219)	2.818(.415)
English	5.624(.311)	3.045(.263)	5.384(.175)	3.113(.263)
	Word meaning ratings[b]			
Chinese	5.775(.223)		1.970(.401)	
English	5.624(.311)		1.888(.432)	

Note. Standard deviations are given in parentheses. Word meanings and vocal tone were each rated on a scale from 1 (*very unpleasant*) to 7 (*very pleasant*).

[a] An analysis of variance (ANOVA) performed on these means showed that only the main effect of vocal tone was significant, $F(1, 56) = 1061.316$, $p < .001$. There was no significant difference in the vocal tone ratings between two languages or between the two types of word meanings.

[b] An ANOVA performed on these means showed that only the main effect of word meaning was significant, $F(1, 28) = 900.274$, $p < .001$. There was no significant difference in word meaning ratings between two languages.

CHAPTER 3

RESULTS

In this chapter I first reported the results regarding accuracy among children in the three cultural groups. Then the results regarding response time were presented.

Preliminary analyses

We first calculated the accuracy for each utterance used in the two tasks across all participants who listened to that utterance. Two words with lower than 0.5 accuracy in the Meaning task were deleted from the following analyses. One word is "Hungry" in English, and the other one is "坏了" (broken) in Chinese. The average accuracies for the rest of utterances across the two tasks were .844 (range: .603-.967), .887 (range: .683-.983), and .858 (range: .623-.977) for EA, CU, and CR sample respectively.

Accuracy

Accuracy in the two tasks

The percentage of accurate responses in each task for each participant was calculated to indicate the accuracy of the judgment. We conducted a 3 (culture: EA vs. CU vs. CR) × 2 (task: Meaning vs. Tone game) × 2 (word meaning: pleasant vs. unpleasant meaning) × 2 (vocal tone: pleasant vs. unpleasant tone) × 2 (gender: boys vs. girls) repeated measure analysis of covariance (repeated measure ANCOVA) with age as a covariate to examine the general pattern of accuracy, with culture and gender as between-subjects variables, and task, meaning, vocal tone as within-subjects variables.

There was a significant main effect of age, whereby older children had higher accuracy than younger children, $F(1, 362) = 188.694$, $p < .001$, $\eta_p^2 = .343$. The main effect of gender was also significant, $F(1, 361) = 7.963$, $p = .005$, $\eta_p^2 = .022$, which was qualified by a marginally significant interaction effect of Gender × Task, $F(1, 361) = 3.786$, $p = .052$, $\eta_p^2 = .010$. Girls tended to have higher accuracy than boys only in the Tone task, $F(1, 369) = 7.473$, $p = .007$, $\eta_p^2 = .020$, but not in the Meaning task, $F(1, 369) = 1.968$, $p = .162$, $\eta_p^2 = .005$. The interaction effect of Task × Gender was presented in Figure 1.

The results revealed a main effect of task, $F(1, 362) = 58.787$, $p < .001$, $\eta_p^2 = .140$, qualified by a marginally significant interaction effect of Task × Culture, $F(1, 362) = 2.685$, $p = .070$, $\eta_p^2 = .015$. In the Meaning task, children from the three cultural groups had the same accuracies, $F(2, 368) = .611$, $p = .543$, $\eta_p^2 = .003$. However, in the Tone task, the effect of culture was significant, $F(2, 368) = 5.537$, $p = .004$, $\eta_p^2 = .029$. Specifically, CU children had significantly higher accuracies than EA children in the Tone task (Bonferroni post hoc tests, $p = .003$; LSD post hoc tests, $p = .001$; Tukey HSD post hoc tests, $p = .003$), and CR children were in the middle, tended to have higher accuracy than EA children (Bonferroni post hoc tests, $p = .200$; LSD post hoc tests, $p = .067$; Tukey HSD post hoc tests, $p = .158$), but were not different from CU children (Bonferroni post hoc tests, $p = .368$; LSD post hoc tests, $p = .123$; Tukey HSD post hoc tests, $p = .270$). By further exploring the interaction effect of Task × Culture, we found that the accuracies in the Meaning task were greater than in the Tone task for all three

cultural groups: EA children, $F(1, 120) = 27.549$, $p < .001$, $\eta_p^2 = .187$, CU children, $F(1, 119) = 5.330$, $p = .023$, $\eta_p^2 = .043$, and CR children, $F(1, 129) = 16.798$, $p < .001$, $\eta_p^2 = .115$. However, the difference in accuracies between the two tasks for EA children was significantly greater than CU children, $F(1, 239) = 11.936$, $p = .001$, $\eta_p^2 = .048$, and CR children, $F(1, 249) = 7.762$, $p = .006$, $\eta_p^2 = .030$. The differences between the two tasks were the same for CU and CR children, $F(1, 248) = .898$, $p = .344$, $\eta_p^2 = .004$. The interaction effect of Task × Culture was presented in Figure 2 and Figure 3.

The main effect of meaning was significant, $F(1, 362) = 12.284$, $p = .001$, $\eta_p^2 = .033$, qualified by an interaction effect of Meaning × Age, $F(1, 362) = 6.205$, $p = .013$, $\eta_p^2 = .017$, and a marginally significant interaction effect of Meaning × Culture, $F(1, 362) = 2.604$, $p = .075$, $\eta_p^2 = .014$. The accuracies when judging pleasant word meaning were higher than unpleasant word meaning for EA, $F(1, 188) = 19.148$, $p < .001$, $\eta_p^2 = .140$, and CR children, $F(1, 127) = 6.391$, $p = .013$, $\eta_p^2 = .048$, but this difference was not significant for CU children, $F(1, 117) = 1.572$, $p = .212$, $\eta_p^2 = .013$. A similar main effect of tone, $F(1, 362) = 5.870$, $p = .016$, $\eta_p^2 = .016$, and an interaction effect of Tone × Culture, $F(2, 362) = 4.117$, $p = .017$, $\eta_p^2 = .022$, proved significant. The advantage of positive tone than negative tone was significant for EA children, $F(1, 118) = 5.359$, $p = .022$, $\eta_p^2 = .043$, and CR children, $F(1, 127) = 9.660$, $p = .002$, $\eta_p^2 = .071$, but not significant for CU children, $F(1, 117) = .951$, $p = .331$, $\eta_p^2 = .008$.

As predicted, there was a significant interaction effect of Meaning × Tone, $F(1, 362) = 61.425$, $p < .001$, $\eta_p^2 = .145$. Judgment accuracies were higher for congruous utterances than incongruous utterances. Further, this interaction was qualified by both task, $F(1, 362) = 36.081$, $p < .001$, $\eta_p^2 = .091$, and culture, $F(2, 362) = 3.113$, $p = .046$, $\eta_p^2 = .017$. The four-way interaction, Task × Meaning × Tone × Culture, proved significant, $F(2, 362) = 4.448$, $p = .012$, $\eta_p^2 = .024$. Interestingly, it also revealed a marginally significant five-way interaction, Task × Meaning × Tone × Culture × Age, $F(2, 362) = 2.491$, $p = .084$, $\eta_p^2 = .014$. Further analyses on the 3-way, 4-way, and 5-way interaction effects were presented in the next section.

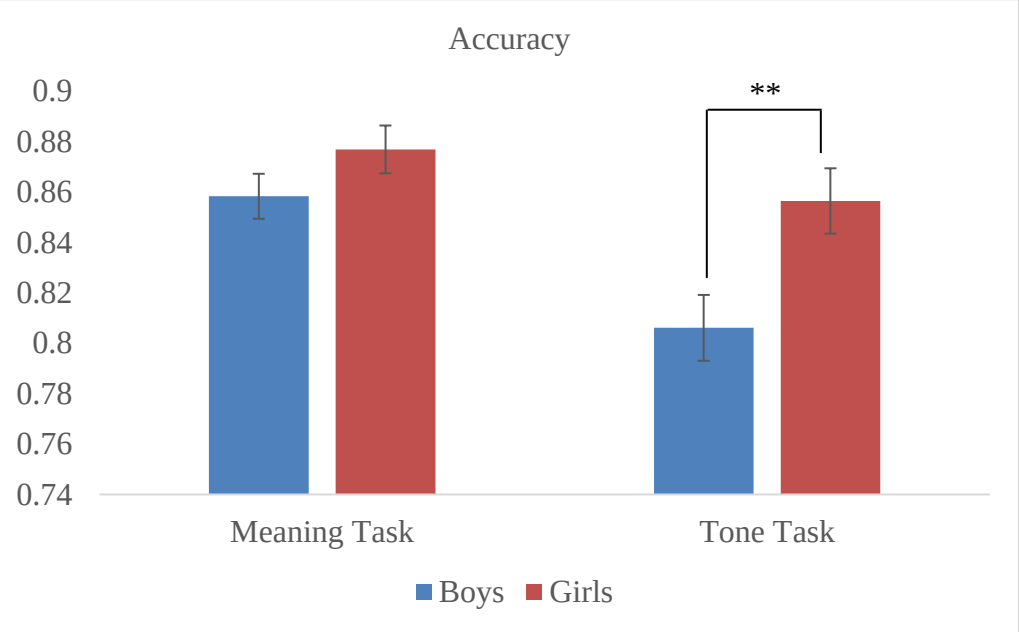


Figure 1. Mean Accuracies as a function of gender and task.

Note. $^{**}p < .01$

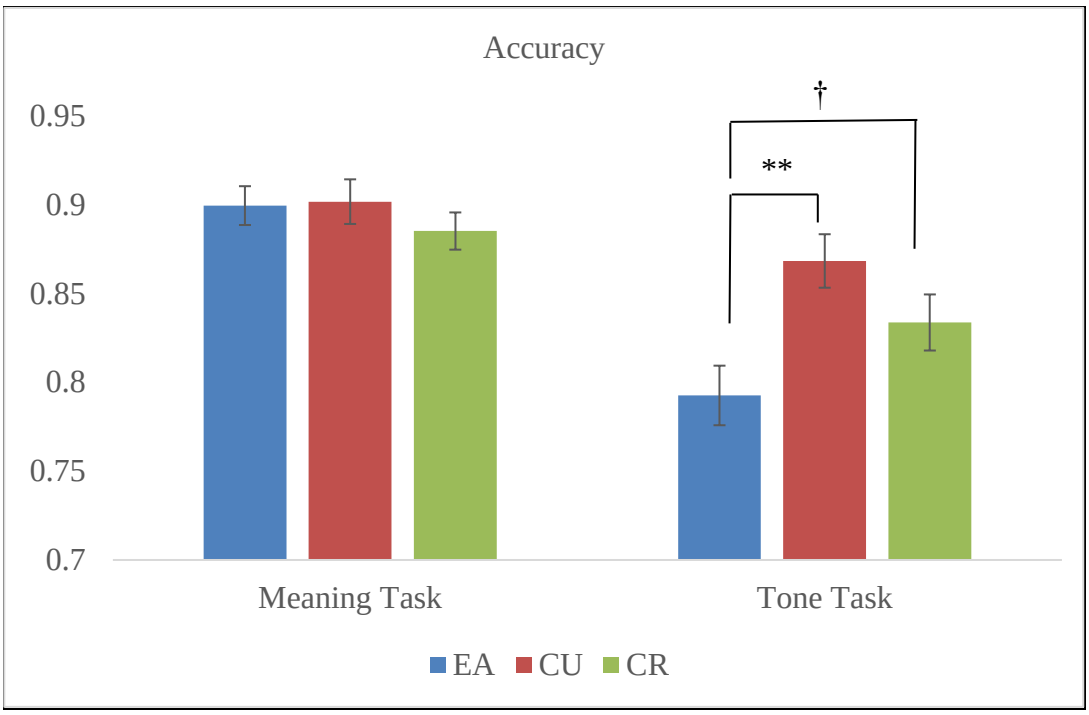


Figure 2. Mean Accuracies as a function of culture and task

Note. $^{**}p < .01$, † $p < .1$

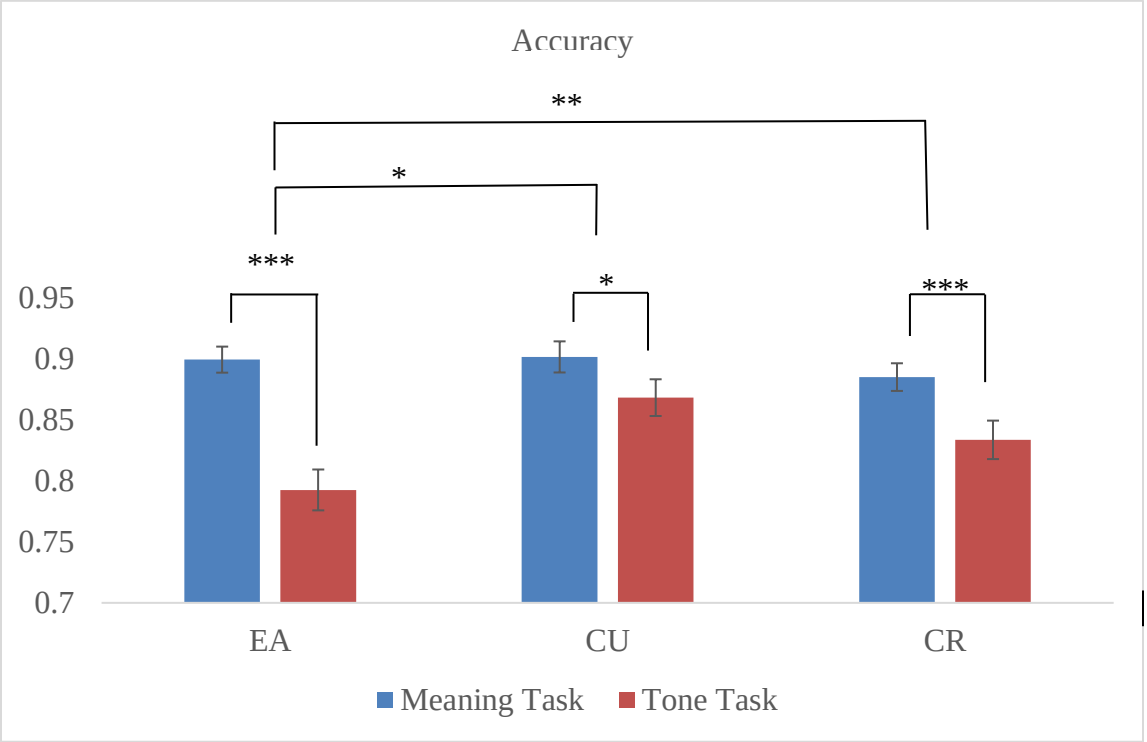


Figure 3. Mean Accuracies in the two tasks as a function of cultural groups

Note. ***$p < .001$ ** $p < .01$ * $p < .05$

Interference effect on accuracy

To facilitate further analyses, we computed an interference index for each participant in each task by subtracting the percentage of correct response for the incongruous utterances (i.e., pleasant word meaning with unpleasant vocal tone; unpleasant word meaning with pleasant vocal tone) from the percentage of correct responses for the congruous utterances (i.e., pleasant in both word meaning and vocal tone; unpleasant in both word meaning and vocal tone). Positive scores indicate the interference by the information in the to-be-ignored aspect of the utterance. For example, the interference score in the Meaning task indicates the interference by the vocal tone, whereas the interference score in the Tone task indicates the interference by the word meaning. We conducted a 3 (culture) × 2 (task) × 2 (gender) repeated measure ANCOVA with age as covariate on the interference index to investigate the interference effects of vocal tone and word meaning in European American and the two groups of Chinese participants. The main effect of culture was significant, $F(2, 362) = 3.159$, $p = .044$, $\eta_p^2 = .017$, whereby EA children's accuracies in general tended to be interfered by the to-be-ignored aspect of utterances to a greater extent than CU children (Bonferroni post hoc tests, $p = .064$; Tukey HSD test, $p = .055$; LSD test, $p = .021$). CR children's interference scores were in the middle but not significantly different from EA or CU children (Bonferroni post hoc tests, $p > .212$). There was also a main effect of age, $F(1, 362) =$

22.234, $p < .001$, $\eta_p^2 = .058$, whereby younger children had a higher level of interference effects than older children.

In terms of the within-subjects effects, there was a main effect of task, $F(1, 362) = 36.077$, $p < .001$, $\eta_p^2 = .091$. In general, the interference scores were higher in the Tone task than in the Meaning task. The interaction effect of Culture × Task proved significant, $F(2, 362) = 4.525$, $p = .011$, $\eta_p^2 = .024$. Separate ANOVAs showed that the interference was greater for Tone task (the influence by word meaning) than for Meaning task (the influence by vocal tone) for EA children, $F(1, 120) = 24.986$, $p < .001$, $\eta_p^2 = .172$, and also CR children, $F(1, 129) = 4.176$, $p = .043$, $\eta_p^2 = .031$. However, for CU children, there was no significant difference in the interference between the two tasks, $F(1, 119) = 2.156$, $p = .145$, $\eta_p^2 = .018$. The difference in the interference scores between two tasks was greater for EA children than for CR children, $F(1, 249) = 7.698$, $p = .006$, $\eta_p^2 = .100$. To further examine the interaction effect of Task × Culture, we conducted ANOVAs on Meaning task and Tone task separately. In the Meaning task, the effect of culture was marginally significant, $F(2, 368) = 2.673$, $p = .070$, $\eta_p^2 = .014$. Specifically, CR children were interfered by vocal tone to a greater extent than EA children (Bonferroni post hoc tests, $p = .070$; LSD post hoc tests, $p = .023$; Tukey HSD post hoc tests, $p = .060$). CU children's interference score in the Meaning task was in the middle, but not different from scores of EA and CR children (Bonferroni post hoc tests, $p > .457$; LSD post hoc tests, $p > .152$; Tukey HSD post hoc tests, $p > .324$). In the Tone task, the effect of

culture on the interference score of accuracy was significant, $F(2, 368) = 5.222$, $p = .006$, $\eta_p^2 = .028$. EA children were influenced more by the word meaning than CU children (Bonferroni post hoc tests, $p = .004$; LSD post hoc tests, $p = .001$; Tukey HSD post hoc tests, $p = .004$) and CR children (Bonferroni post hoc tests, $p = .139$; LSD post hoc tests, $p = .046$; Tukey HSD post hoc tests, $p = .114$). CU and CR children's interference scores in the Tone task were not significantly different (Bonferroni post hoc tests, $p = .622$; LSD post hoc tests, $p = .207$; Tukey HSD post hoc tests, $p = .417$). The mean interference scores for accuracy by culture and task were presented in Figure 4 and Figure 5.

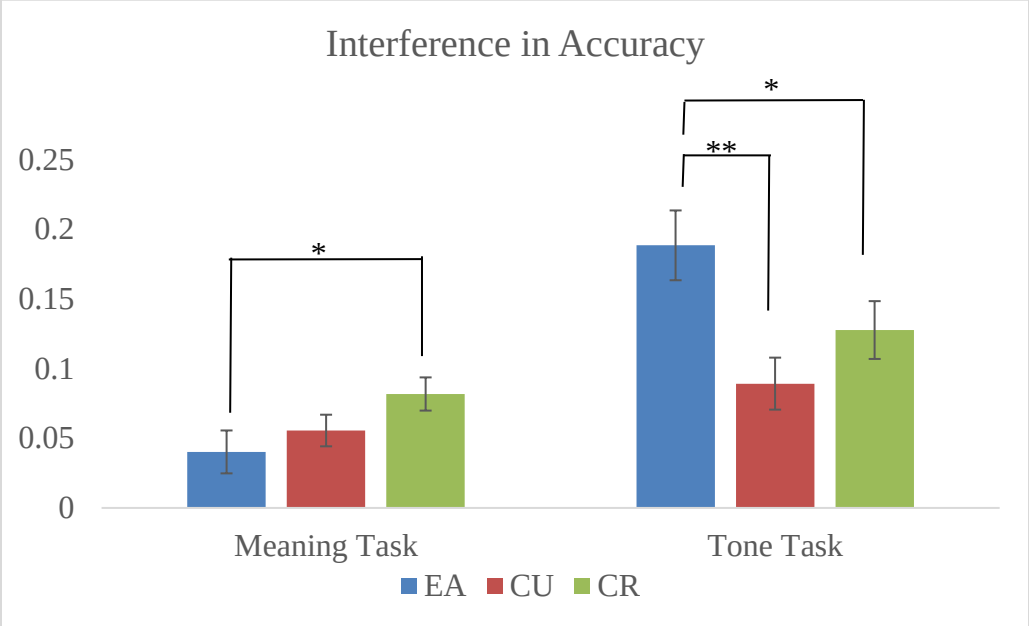


Figure 4. Interference score in accuracy as a function of culture and task

Note. The symbols denote the significance level of the differences based on LSD post hoc tests. ** $p < .01$ * $p < .05$

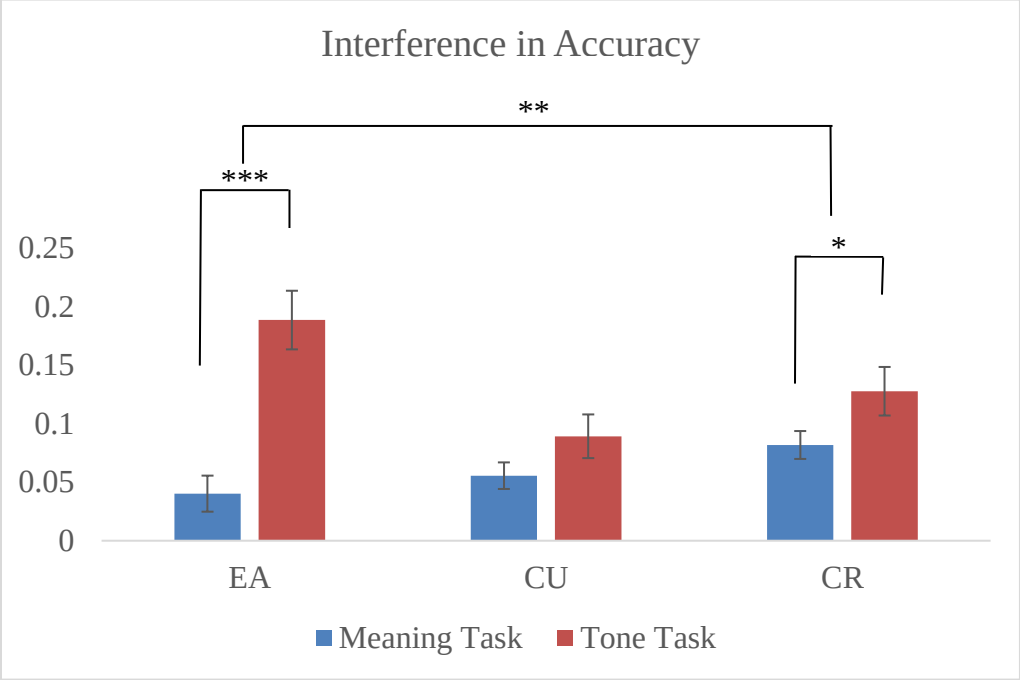


Figure 5. Interference score in accuracy as a function of task and culture

Note. ***$p < .001$ * $p < .05$

Moreover, the 3-way interaction effect Task × Culture × Age was marginally significant, $F(2, 362) = 2.548$, $p = .080$, $\eta_p^2 = .014$. In order to further understand this 3-way interaction effect, we conducted ANOVAs on interference scores by 3 age groups (4-5, 6-7, and 8-9-year-old) separately. Results showed that for the two younger groups (age 4-5 and age 6-7), there was no cultural difference in the interference scores in the Meaning task, $F = .352$ - $.716$, $p = .491$ - $.704$, $\eta_p^2 = .006$ - $.012$. However, for older children (age 8-9), the effect of culture on the interference scores in the Meaning task was

significant, $F(2, 123) = 6.295$, $p = .002$, $\eta_p^2 = .093$. Specifically, EA children experienced less interference effect in the Meaning task than CU children (Bonferroni post hoc tests, $p = .061$; LSD post hoc tests, $p = .020$; Tukey HSD post hoc tests, $p = .053$), and CR children (Bonferroni post hoc tests, $p = .002$; LSD post hoc tests, $p = .001$; Tukey HSD post hoc tests, $p = .002$). CU and CR children experienced the same level of interference scores in the Meaning task (Bonferroni post hoc tests, $p = .826$; LSD post hoc tests, $p = .275$; Tukey HSD post hoc tests, $p = .519$). This suggested that Chinese children were interfered by vocal the tone more than EA children only for 8-9-year-olds.

In terms of the interference score of accuracy for the Tone task, there were significant cultural differences in the interference scores for 4-5 year-olds, $F(2, 120) = 4.120$, $p = .019$, $\eta_p^2 = .064$. For these youngest children, EA children had higher interference scores than CU children (Bonferroni post hoc tests, $p = .025$; LSD post hoc tests, $p = .008$; Tukey HSD post hoc tests, $p = .022$) and CR children (Bonferroni post hoc tests, $p = .086$; LSD post hoc tests, $p = .029$; Tukey HSD post hoc tests, $p = .073$). There was no difference in the interference scores in the Tone task between the youngest CU and CR children (Bonferroni post hoc tests, $p = .1.000$; LSD post hoc tests, $p = .613$; Tukey HSD post hoc tests, $p = .868$). The effect of culture was marginally significant for 6-7 year-olds, $F(2, 119) = 2.453$, $p = .090$, $\eta_p^2 = .040$. Bonferroni post hot tests showed no significant difference between any two cultural groups ($p > .142$). LSD post doc tests revealed that EA children had higher interference scores than CU children in the Tone

task (p = .047), but had the same level of interference scores in the Tone task as CR children (p = .842). CR children had higher interference scores than CU children, but the difference was marginally significant (p = .068). However, older children (age 8-9) in the three cultural groups experienced the same level of interference effects in the Tone task, $F(2, 123) = .121$, $p = .886$, $\eta_p^2 = .002$. The mean interference scores by culture and age group are displayed in Figure 6, Figure 7, and Figure 8.

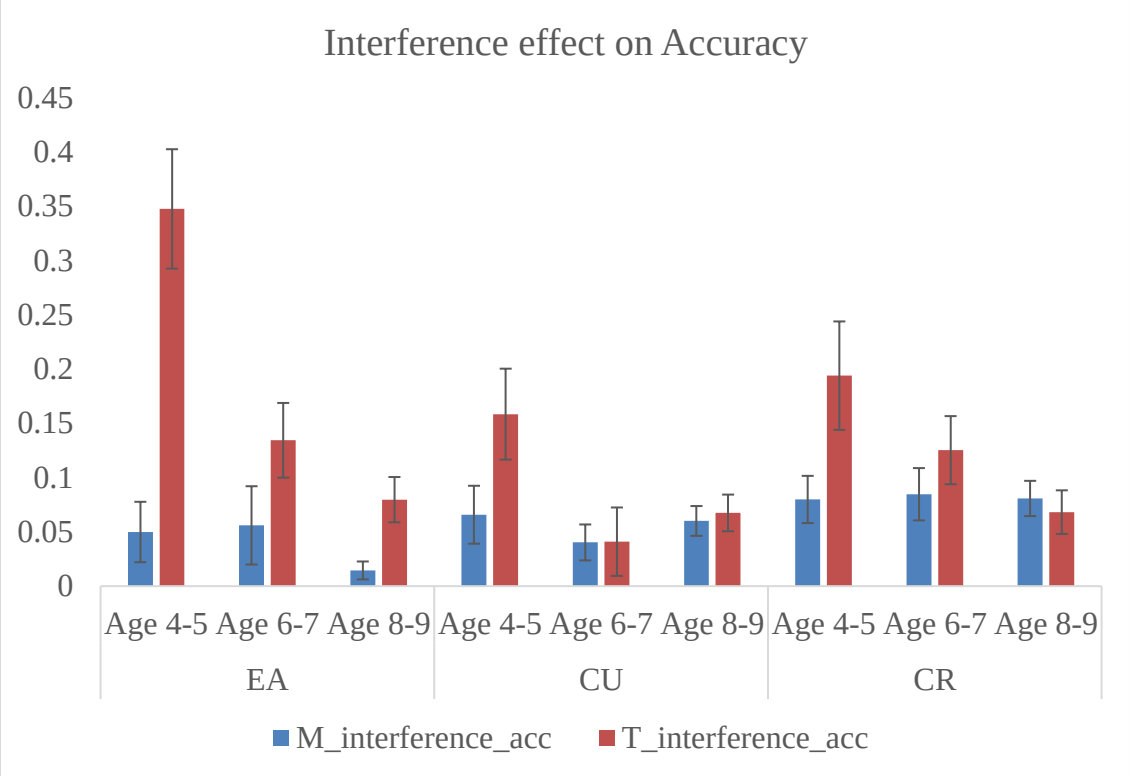


Figure 6. Interference scores with standard errors as a function of age group, culture, and task.

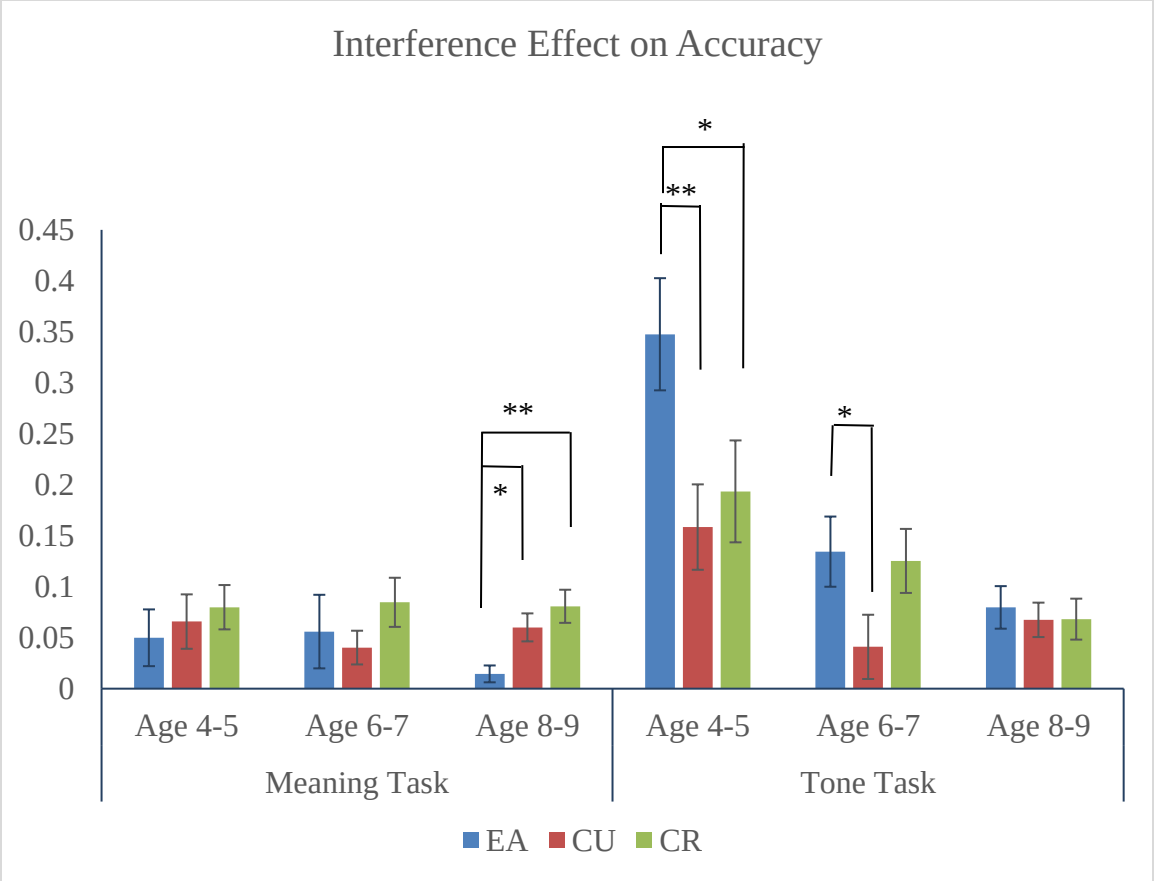


Figure 7. Interference scores with standard errors as a function of age group, task, and culture.

Note. The symbols denote the significance level of the differences based on LSD post hoc tests ** $p < .01$ * $p < .05$

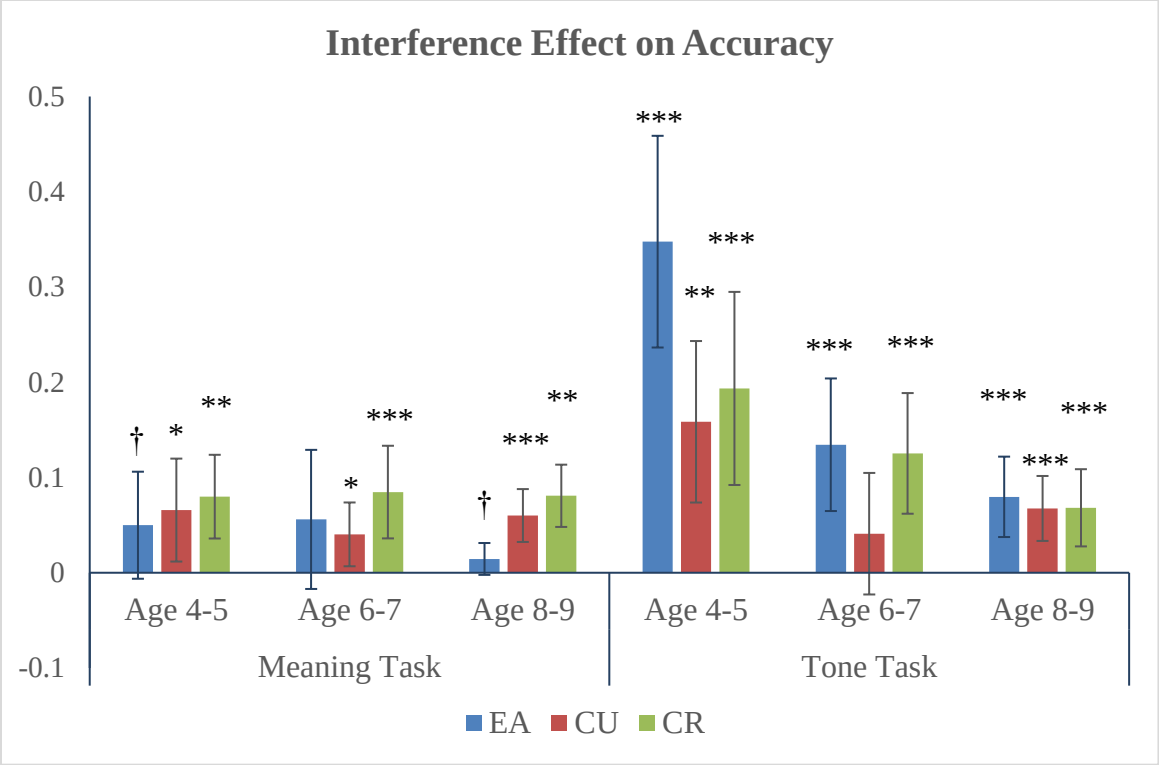


Figure 8. Interference scores with 95% confidence intervals as a function of age group, task, and culture.

Note. The symbols denote whether the interference scores were different from zero.

*** $p < .001$ ** $p < .01$ * $p < .05$ † $p < .1$

Response time

Response time in the two tasks

For correct responses, the response time was analyzed in the same way as analyzing accuracy except for computing the interference index. The interference scores of response time were calculated by subtracting the mean response time of the congruous utterances from the mean response time of the incongruous utterances. Therefore, positive scores still indicate an interference effect of the to-be-ignored aspect of the utterance.

We submitted data on the response time (ms) to a repeated measure ANCOVA. There was a significant main effect of age, whereby older children had shorter response time than younger children, $F(1, 335) = 166.452$, $p < .001$, $\eta_p^2 = .332$. The main effect of culture was also significant, $F(2, 335) = 4.990$, $p = .007$, $\eta_p^2 = .029$, which qualified by an interaction effect of Culture × Age, $F(2, 335) = 4.337$, $p = .014$, $\eta_p^2 = .025$. In order to understand the interaction effect between age and culture, we examined cultural differences in response time for the 3 different age groups. The effects of culture on response time were marginally significant for 4-5-year-olds, $F(2, 120) = 3.057$, $p = .051$, $\eta_p^2 = .048$, significant for 8-9 year-olds, $F(2, 123) = 3.174$, $p = .045$, $\eta_p^2 = .049$, but not significant for 6-7 year-olds, $F(2, 119) = .252$, $p = .777$, $\eta_p^2 = .004$. Specifically, for 4-5-year-old children, EA children had shorter response time than CR children (Bonferroni post hoc tests, $p = .047$, LSD post hoc tests, $p = .016$, Tukey HSD post hoc tests, p

= .041). CU children's response time was the same as EA and CR children (Bonferroni post hoc tests, p = .435 – 1.000, LSD post hoc tests, p = .145 - .336, Tukey HSD post hoc tests, p = .310 - .600). For 8-9-year-olds, CU children had longer response time than CR children (Bonferroni post hoc tests, p = .065, LSD post hoc tests, p = .022, Tukey HSD post hoc tests, p = .056) and EA children (Bonferroni post hoc tests, p = .173, LSD post hoc tests, p = .046, Tukey HSD post hoc tests, p = .112). EA children's response time was not different from CR children (Bonferroni post hoc tests, p = 1.000, LSD post hoc tests, p = .806, Tukey HSD post hoc tests, p = .967).

In terms of within-subjects effects, there was a significant main effect of task, $F(1, 335) = 20.446$, $p < .001$, $\eta_p^2 = .058$, with shorter response time in the word meaning task compared to the vocal tone task. It was qualified by an interaction effect of Task × Age, $F(1, 335) = 20.192$, $p < .001$, $\eta_p^2 = .057$. We further looked at the effects of task in different age groups. The response time in the Meaning task tended to be shorter than in the Tone task for 4-5-year-old children, $F(1, 120) = 3.042$, $p = .084$, $\eta_p^2 = .025$, but the response time was longer in the Meaning task than in the Tone task for 8-9-year-old children, $F(1, 123) = 10.749$, $p = .001$, $\eta_p^2 = .080$. For 6-7-year-olds, there was no difference in the response time between two tasks, $F(1, 119) = .915$, $p = .341$, $\eta_p^2 = .008$. There was also a significant interaction effect of Task × Culture, $F(2, 335) = 6.688$, $p = .001$, $\eta_p^2 = .038$. In the Meaning task, the effect of culture was marginally significant, $F(2, 368) = 2.716$, $p = .067$, $\eta_p^2 = .015$. Specifically, in the Meaning task EA children

tended to have shorter response time than CU children (Bonferroni post hoc tests, p = .101; LSD post hoc tests, p = .034; Tukey HSD post hoc tests, p = .085) and CR children (Bonferroni post hoc tests, p = .177; LSD post hoc tests, p = .059; Tukey HSD post hoc tests, p = .142). There was no difference in response time between CR and CU children (Bonferroni post hoc tests, p = 1.000; LSD post hoc tests, p = .779; Tukey HSD post hoc tests, p = .957). In the Tone task, there was no significant cultural difference in the response time, $F(2, 368) = .817$, $p = .443$, $\eta_p^2 = .004$. To further explore the interaction effect of Task × Culture, we examined the differences in response time between the two tasks for each cultural group. The difference was not significant for any of the three cultural groups, $F = .095 - .178$, $p = .178 - .759$, $\eta_p^2 = .001 - .015$. The interaction effect of Task × Culture was presented in Figure 9 and Figure 10. The 3-way interaction Age × Culture × Task was also significant, $F(2, 335) = 6.207$, $p = .002$, $\eta_p^2 = .036$.

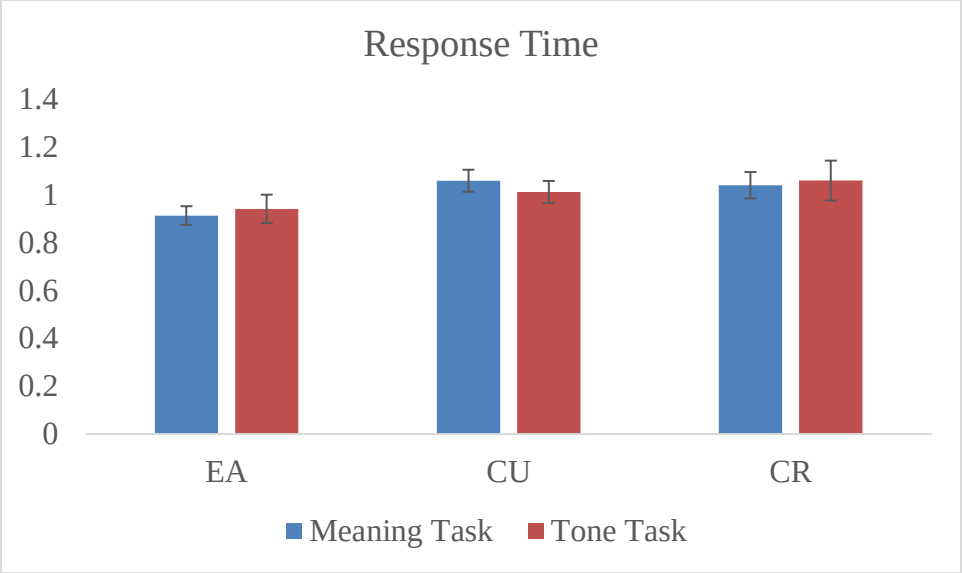


Figure 9. Response time in the two tasks as a function of culture and task.

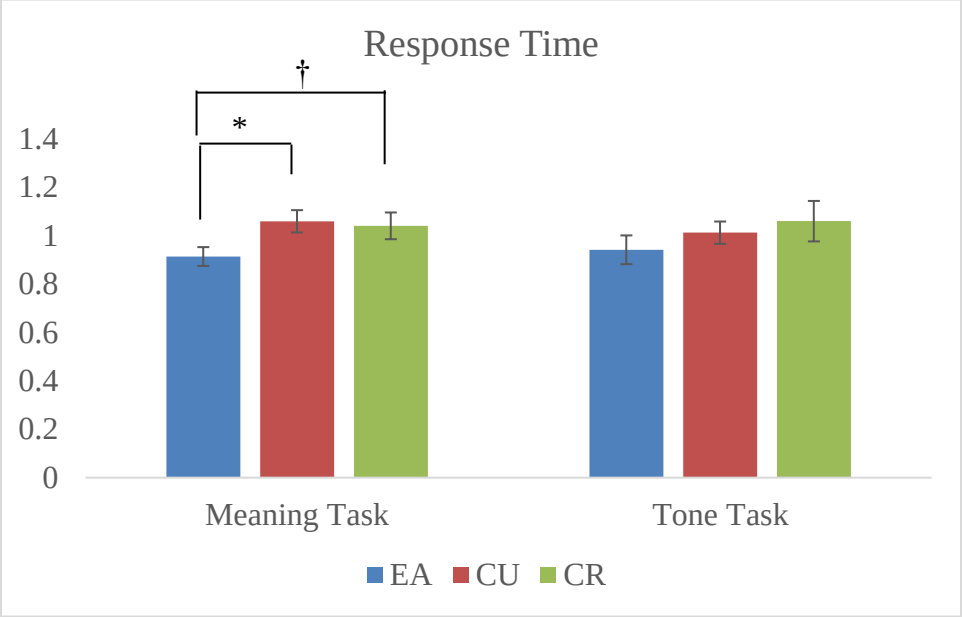


Figure 10. Response time in the two tasks as a function of task and culture.

Note. * $p < .05$ † $p < .1$

The main effect of meaning was not significant, but the interaction effect of Meaning × Culture was marginally significant, $F(1, 335) = 2.943$, $p = .054$, $\eta_p^2 = .017$. Specifically, the response time was longer for the utterances with a negative meaning than for utterances with a positive meaning, only for CR children, $F(1, 118) = 4.114$, $p = .045$, $\eta_p^2 = .034$, but not for EA, $F(1, 103) = 1.481$, $p = .226$, $\eta_p^2 = .014$ or CU children, $F(1, 112) = 1.043$, $p = .309$, $\eta_p^2 = .009$. There was a main effect of tone, $F(1, 335) = 5.339$, $p = .021$, $\eta_p^2 = .016$. In general, the response time was longer for the utterances with a negative tone than for utterances with a positive tone. The interaction effect of Meaning × Tone was significant, $F(1, 335) = 20.313$, $p < .001$, $\eta_p^2 = .057$. Response time was shorter for congruent utterances than for incongruent utterances. Further, this interference effect was more pronounced in the Tone task, $F(1, 335) = 5.236$, $p = .023$, $\eta_p^2 = .015$, and for younger children, $F(1, 335) = 5.160$, $p = .024$, $\eta_p^2 = .015$, than in the Meaning task and for older children respectively. The four-way interaction effects of Meaning × Tone × Task × Culture, $F(2, 335) = 3.621$, $p = .028$, $\eta_p^2 = .021$, and Meaning × Tone × Task × Age, $F(1, 335) = 4.497$, $p = .035$, $\eta_p^2 = .013$, were significant. The five-way interaction effect of Meaning × Tone × Task × Culture × Age was marginally significant, $F(2, 335) = 2.705$, $p = .068$, $\eta_p^2 = .016$.

The interference effect on response time

Further analyses on the interference effects on response time confirmed the main effect of age on the interference, $F(1, 356) = 7.263$, $p = .007$, $\eta_p^2 = .020$. It showed that

the interference effects on response time decreased with age. The main effect of culture on the interference scores also proved significant, $F(2, 356) = 4.236$, $p = .015$, $\eta_p^2 = .023$. In general, EA children had lower interference scores in response time than CU children (Bonferroni post hoc tests, $p = .005$; LSD post hoc tests, $p = .002$; Tukey HSD post hoc tests, $p = .005$), and CR children (Bonferroni post hoc tests, $p = .001$; LSD post hoc tests, $p < .001$; Tukey HSD post hoc tests, $p = .001$). There was no difference in the interference scores of the response time between CU and CR children (Bonferroni post hoc tests, $p = 1.000$; LSD post hoc tests, $p = .789$; Tukey HSD post hoc tests, $p = .961$). The interaction effect of Task × Age was significant, $F(1, 356) = 4.707$, $p = .031$, $\eta_p^2 = .013$. In the Meaning task, the interference scores were not significantly correlated with age, $r = -.015$, $p = .769$, but significantly decreased with age in the Tone task, $r = -.139$, $p = .008$.

The results revealed a significant interaction effect of Culture × Task, $F(2, 356) = 5.415$, $p = .005$, $\eta_p^2 = .030$. Separate ANOVAs showed that the effects of culture on the interference scores of response time were significant in both Meaning task, $F(2, 366) = 9.667$, $p < .001$, $\eta_p^2 = .050$, and Tone task, $F(2, 364) = 3.363$, $p = .036$, $\eta_p^2 = .018$. Specifically, in the Meaning task, EA children's interference scores were lower than CU children (Bonferroni post hoc tests, $p < .001$; LSD post hoc tests, $p < .001$; Tukey HSD post hoc tests, $p < .001$) and CR children (Bonferroni post hoc tests, $p = .021$; LSD post hoc tests, $p = .007$; Tukey HSD post hoc tests, $p = .019$). CU children tended to have

lower interference scores than CR children (Bonferroni post hoc tests, $p = .260$; LSD post hoc tests, $p = .087$; Tukey HSD post hoc tests, $p = .200$). In the Tone task, EA children had lower interference scores than CR children (Bonferroni post hoc tests, $p = .030$; LSD post hoc tests, $p = .010$; Tukey HSD post hoc tests, $p = .027$). There were no differences between CU children and EA or CR children (Bonferroni post hoc tests, $p = .462$ - $.784$; LSD post hoc tests, $p = .154$ - 261; Tukey HSD post hoc tests, $p = .327$ - $.499$). The interaction effect of Culture × Task was examined from a different angle by looking at the effect of task for EA, CU, and CR groups separately. The interference scores in the Meaning task were the same as the scores in the Tone task for EA children, $F(1, 118) = .945$, $p = .333$, $\eta_p^2 = .008$, and CU children, $F(1, 116) = 1.114$, $p = .293$, $\eta_p^2 = .010$. However, for CR children, the interference scores in the Tone task tended to be higher than in the Meaning task, $F(1, 128) = 3.876$, $p = .051$, $\eta_p^2 = .029$. The interference scores of response time by culture and task are presented in Figure 11 and Figure 12.

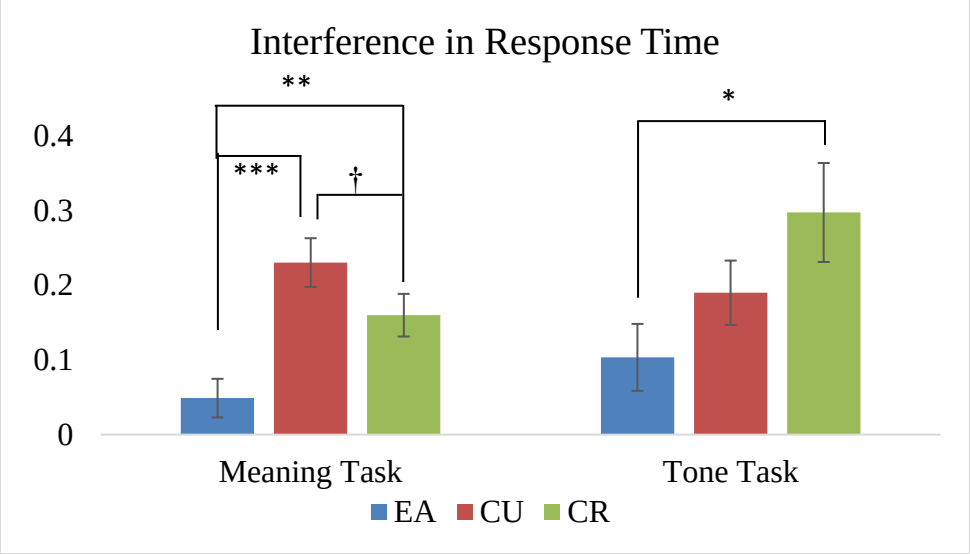


Figure 11. Interference effects on response time as a function of task and culture.

Note. *** $p < .001$ ** $p < .01$ * $p < .05$ † $p < .1$

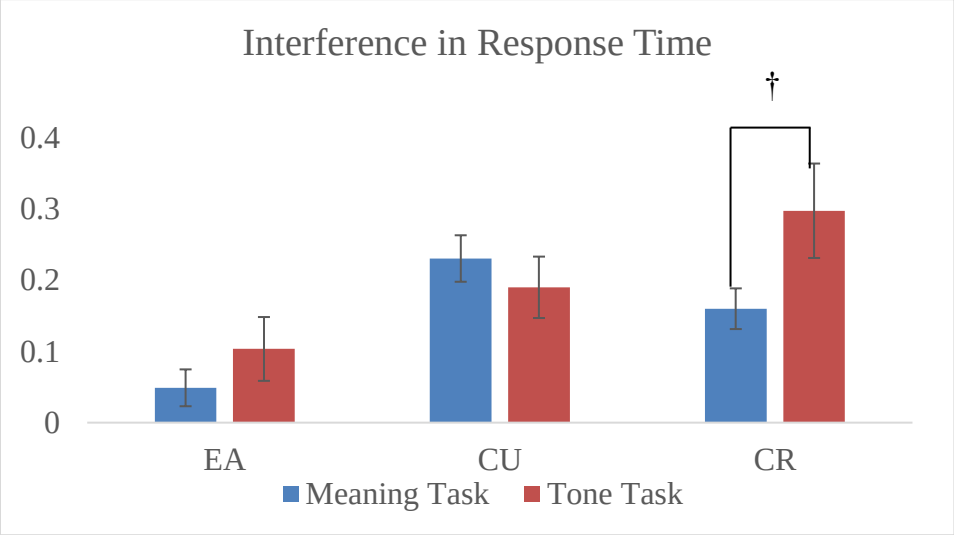


Figure 12. Interference effects on response time as a function of culture and task.

Note. † $p < .1$

The three-way interaction effect of Culture × Task × Age also proved significant, $F(2, 356) = 4.048$, $p = .018$, $\eta_p^2 = .022$. To further analyze the interaction effect, we looked at the cultural different in the two tasks for the three age groups separately. In the Meaning task, the effect of culture was significant for 4-5-year-old children, $F(2, 118) = 6.920$, $p = .001$, $\eta_p^2 = .105$, and 8-9-year-olds, $F(2, 123) = 4.230$, $p = .017$, $\eta_p^2 = .064$, but not for 6-7-year-olds, $F(2, 119) = 1.505$, $p = .226$, $\eta_p^2 = .025$. Specifically, for 4-5-year-olds, EA children were interfered by the vocal tone less than CU children (Bonferroni post hoc tests, $p = .001$; LSD post hoc tests, $p < .001$; Tukey HSD post hoc tests, $p = .001$) and CR children (Bonferroni post hoc tests, $p = .272$; LSD post hoc tests, $p = .091$; Tukey HSD post hoc tests, $p = .207$). CR young children tended to have lower interference scores than CU children (Bonferroni post hoc tests, $p = .133$; LSD post hoc tests, $p = .044$; Tukey HSD post hoc tests, $p = .109$) in terms of the interference scores in the Meaning task. For 8-9-year-olds, EA children were interfered by the vocal tone less than CU children (Bonferroni post hoc tests, $p = .138$; LSD post hoc tests, $p = .046$; Tukey HSD post hoc tests, $p = .112$) and CR children (Bonferroni post hoc tests, $p = .016$; LSD post hoc tests, $p = .005$; Tukey HSD post hoc tests, $p = .015$). There was no difference between CU and CR oldest children (Bonferroni post hoc tests, $p = 1.000$; LSD post hoc tests, $p = .437$; Tukey HSD post hoc tests, $p = .716$). In the Tone task, the effect of culture was significant for 4-5 year old children, $F(2, 117) = 3.167$, $p = .046$, $\eta_p^2 = .051$, but not for 6-7 year-olds, $F(2, 118) = .120$, $p = .887$, $\eta_p^2 = .002$, or 8-9-year-olds,

$F(2, 123) = 1.888$, $p = .156$, $\eta_p^2 = .030$. Specifically, for 4-5-year-olds, CR children had lower interference scores than EA children (Bonferroni post hoc tests, $p = .049$; LSD post hoc tests, $p = .016$; Tukey HSD post hoc tests, $p = .043$), and CU children (Bonferroni post hoc tests, $p = .265$; LSD post hoc tests, $p = .088$; Tukey HSD post hoc tests, $p = .202$). There was no difference between EA and CU children (Bonferroni post hoc tests, $p = 1.000$; LSD post hoc tests, $p = .481$; Tukey HSD post hoc tests, $p = .760$). The mean interference scores by task, culture, and age group are displayed in Figure 13, Figure 14, and Figure 15.

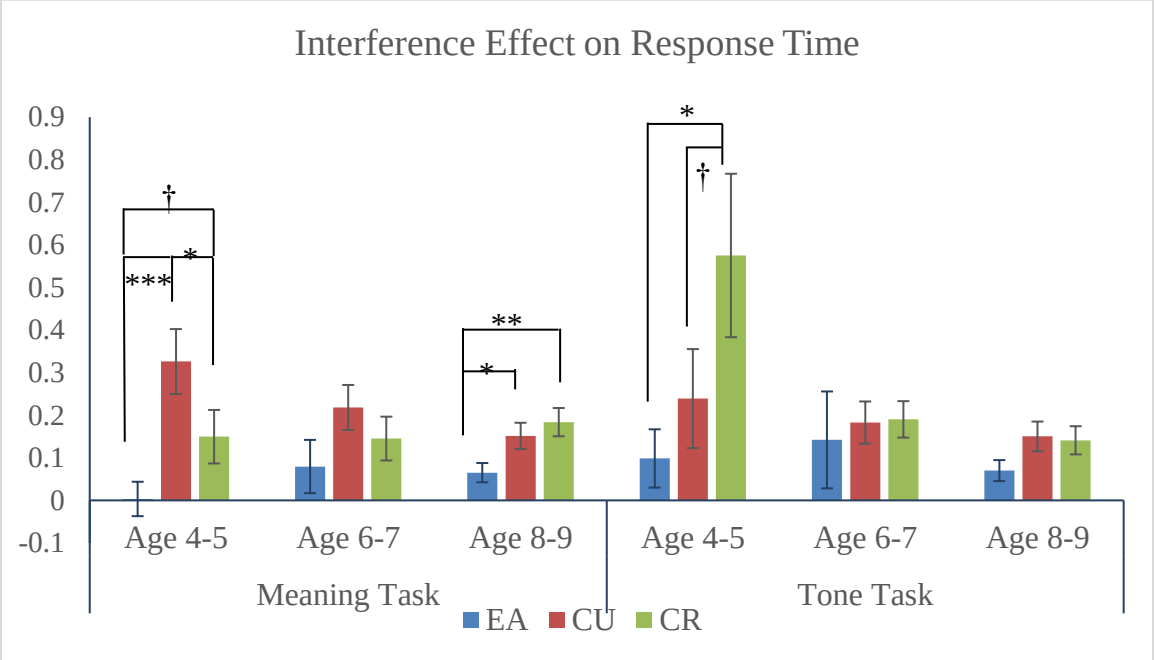


Figure 13. Interference scores of RT with standard error as a function of age group, task, and culture.

Note. *** $p < .001$ ** $p < .01$ * $p < .05$ † $p < .1$

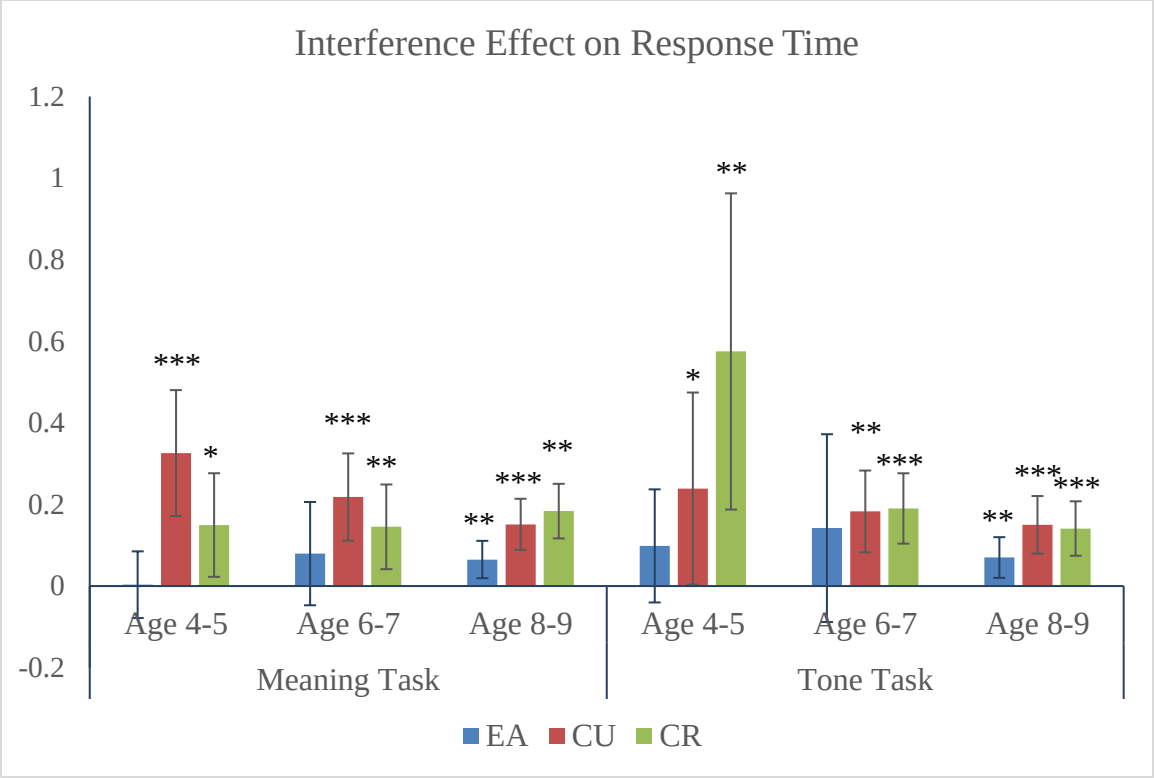


Figure 14. Interference scores of RT with 95% confidence intervals as a function of age group, task, and culture.

Note. The symbols denote whether the interferences scores were different from zero.

*** $p < .001$ ** $p < .01$ * $p < .05$ † $p < .1$

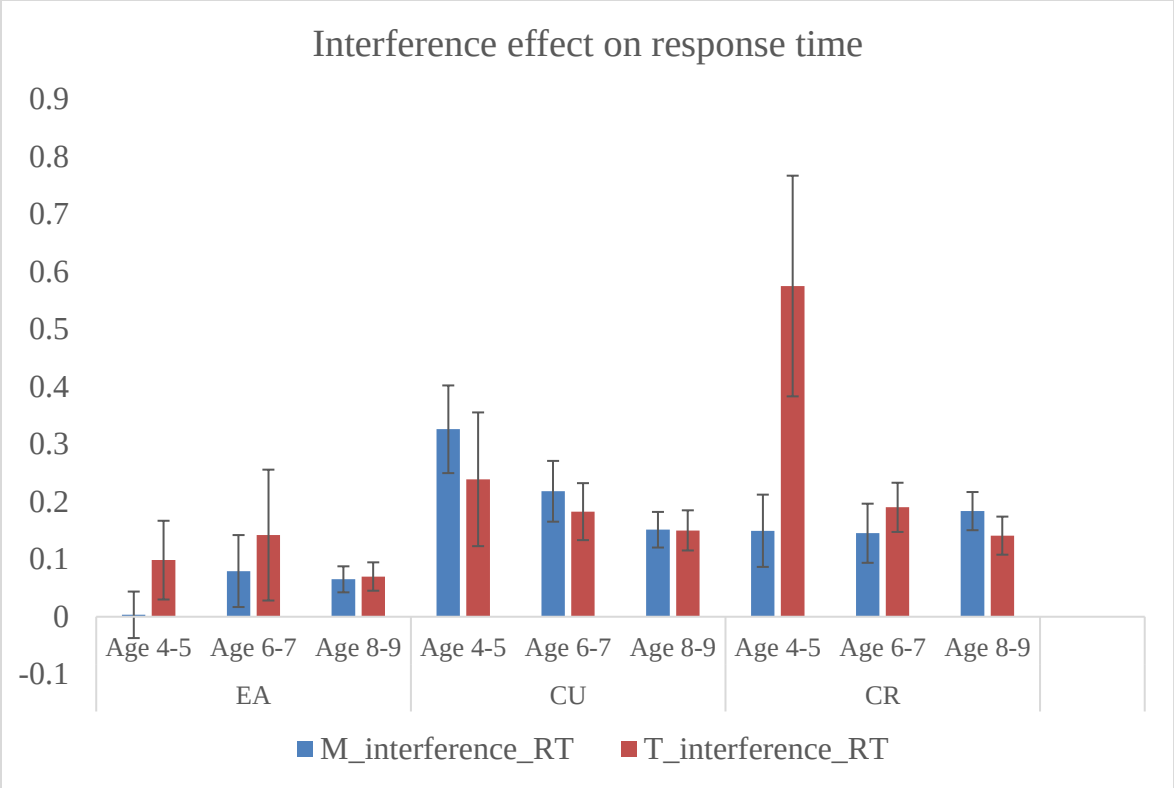


Figure 15. Interference scores of RT with standard errors as a function of age group, culture, and task.

CHAPTER 4

DISCUSSION

This dissertation examined European American, Chinese urban, and Chinese rural children's attention to word meanings and vocal tones while listening to spoken words. Results from the three groups of children have provided insights to cultural similarities and differences in children's sensitivity to different emotional cues from auditory utterances.

Consistent with my hypotheses, there were differences in the patterns of attention to word meanings and vocal tones between European American and Chinese children in China. European American children significantly differed from both Chinese urban and Chinese rural children in the performances in the tasks, nonetheless the two groups of Chinese children performed very similarly to each other. European American children were more sensitive to word meanings and less sensitive to vocal tones than Chinese children. The cultural differences in attention to vocal tones were more pronounced for older children in the age group of 8-9 years old. The cultural differences were revealed in both overall performances of accuracies and response time in the two tasks, as well as the interference effects on the accuracies and response time by the to-be-ignored aspect of emotional utterances. Below I explain the major findings, starting from the overall accuracy and response time performed on the Meaning task and the Tone task, followed by the interference effects in these two tasks.

Overall accuracy and response time

We found significant cultural differences in accuracies in the Tone task, and cultural differences in response time in the Meaning task. Specifically, Chinese children performed more accurately than European American children in the Tone tasks, whereas European American children performed faster in the Meaning task than Chinese children. As we argued in the introduction, people in high-context cultures, such as Chinese culture, are usually encouraged to express and detect emotional cues through contextual information (Ambady, et al., 1996; Hall, 1976). Children in such high-context cultures may be socialized to pay more attention to vocal tones, an important type of contextual information. Thus, Chinese children had higher accuracies when judging vocal tones, compared to European American children. In contrast, in low-context cultures, such as European American culture, people are generally encouraged to explicitly talk about emotions and directly communicate emotions (Hall, 1976; Wang, 2003). As a result, European American children are more sensitive to the more explicit aspect of the utterances: word meanings. Therefore, European American children performed faster when judging wording meanings than Chinese children.

In addition to cultural differences, we also found gender differences in children's accuracies in the Tone task. Girls in general performed better than boys in the Tone task. When judging the vocal tones of the utterances, girls had higher accuracies than boys. This is consistent with previous findings that girls usually have greater emotion

knowledge than boys (Yang & Wang, 2016). Girls' advantage in understanding emotions may be attributed to their more practice on recognizing and expressing emotions, as well as their parents' emphasis on emotions during emotion socialization practices. Studies have shown that girls talk more about feelings than boys (Fivush, Brotman, Buckner, & Goodman, 2000) and parents discuss emotional states more frequently with girls than with boys (Fivush, 1998; Wang, 2001). Girls' greater involvement in emotional practice may help them understand emotions better than boys in general. However, empirical research showed mixed results. Some research revealed no gender differences in emotion knowledge (Doan & Wang, 2010; Wang, 2003). On the other hand, Wang, Hutt, Kulkofsky, McDermott, and Wei (2006) found that Chinese girls had greater emotion knowledge than Chinese boys, but no gender difference was found in European American children. This study further investigated this issue in a different task about implicit emotion knowledge and suggested that girls are better at judging emotional tones than boys.

Noticeably, across all three cultural groups, children had higher accuracies in the Meaning task than in the Tone task. One possible explanation concerns the fact that preschool and middle childhood are periods when children's language develops quickly. When children are consistently learning vocabularies, the sematic content of the utterance may draw their attention more than the vocal tone. It is also possible that word meanings are usually more unequivocal than vocal tones, and the word meaning may be easier to

judge for all children regardless of their cultural backgrounds. Therefore, children in all three cultural groups were able to judge the pleasantness of word meanings more accurately than vocal tones, although the difference between the two tasks was more pronounced among European American children than Chinese children.

We also found the main effects of meaning and tone, whereby, children judged the pleasant words and pleasant tones more accurately and quickly compared to unpleasant words and unpleasant tones respectively. This is consistent with findings from studies on children's emotion recognition of facial expressions. Children recognize positive facial expressions more accurately and faster than negative emotional facial expressions (De Sonneville, Vershoor, Njiokiktjien, Veld, Op het, Toorenaar, & Vranken, 2002; Markham & Wang, 1996; Widen & Russel, 2010). Previous researchers argued that the positive advantage is primarily due to the features of happy face which are easier to recognize (Markham & Wang, 1996), and the lack of other positive emotions which can compete with happiness (De Sonneville, et al., 2002). However, the results in the present study suggested that there is a positive advantage not only in recognizing facial expressions but also in recognizing emotions from word meanings and vocal tones. The positive advantage is present among both European American and Chinese children regardless of cultural background and socioeconomic status. It is reasonable to argue that the reason for positive advantage should be more robust than the features of facial expression of happiness. One possible explanation is that people are more likely to

express positive emotions towards a child. Children's more frequent perceptions of positive emotions make them become more familiar with positive emotional expressions, pleasant tones, and good words compared to negative ones. As a result, children are able to recognize positive emotions more accurately and quickly than negative emotions from facial expressions, word meanings, and vocal tones.

Interference effects on accuracy and response time

Cultural differences in the interference effects found in this study echoed the results regarding the overall accuracies and response time. European American children were interfered by word meanings more and were interfered by vocal tones less than Chinese children. By analyzing it in another way, we found that word meanings were more interfering than vocal tones for European American children. Conversely, vocal tones seemed to have a larger interference effect on Chinese children compared to word meanings. As discussed above, as members in a higher-context culture, Chinese children are more likely to be encouraged to address vocal tones in their emotional communication. They may be trained to be more sensitive to vocal tones than children from a low-context culture, such as European American culture. Gradually, their attention to vocal tones becomes a relatively automatic process and harder to inhibit. Therefore, it requires more effort for Chinese children to ignore vocal tones than for European American children. In the Meaning task, when asked to judge the pleasantness of word meanings but to ignore the vocal tones, Chinese children were interfered by their

automatic processing of vocal tones to a greater extent than their European American counterparts. In contrast, European American children are usually encouraged to express their emotions with explicit words, so they are socialized to be more familiar with and sensitive to word meanings than Chinese children. Consequently, word meanings were too salient to ignore for European American children, and had a greater interference effect on them, compared to Chinese children.

When looking at accuracy and response time separately, we found discrepancies in the results between these two measurements. In terms of accuracy, European American had higher interference scores than Chinese children, whereas in terms of response time Chinese children had higher interference scores than European American counterparts, across the two tasks. One explanation is that Chinese parents value their children's academic achievement strongly and have high expectations of their performance on tests (Chao, 1996; Campbell & Mandel, 1990). Chinese children may be socialized to put a great emphasis on scores which are often determined by accuracies on the tests. When doing tasks, Chinese children maybe accustomed to focus on accuracies, but sacrifice response time if necessary. Thus, in the tasks of this study, Chinese children might set accuracy as a higher priority, and prolong response time to maintain a high accuracy rate when interfered, compared to European American children who did the opposite.

Noticeably, the greater interference effect on response time for Chinese rural children even emerged in the Tone task, although their interference scores in accuracy in

the Tone task were lower than European American children. That means, inconsistent with our hypothesis, Chinese rural children's response time was interfered by the word meaning more than European American children. However, Chinese rural children's higher interference scores in response time was primarily driven by the long response time of 4-5-year-olds. This youngest group of Chinese rural children had much longer response time than older Chinese rural children or the young children in other groups. It is possible that these young rural children had limited access to similar computer games and they were not familiar with the rules when instructed to response as fast as possible. Additionally, the response time is influenced by children's skills other than their cognitive attention, such as motor skills. Therefore, for young children, accuracy might be a more appropriate index in such tasks.

There were interesting effects of age. In terms of accuracy, the cultural differences in the interference scores in the Meaning task were only significant for children aged 8-9 years, but not for the two younger groups of children, whereas in the Tone task the cultural differences existed in the two younger groups of children. The results suggest that emotion socialization may play a role in the cultural differences in the patterns of attention to vocal tones. As children grow older, their ability of inhibition develops, so their interference scores tend to decline over time. Unlike European American children, Chinese children's interference scores in the Meaning task did not decrease over time. As a result, a significant cultural difference in the sensitivity to vocal

tones emerge for 8-9-year-old children. It is reasonable to argue that Chinese children may be socialized to keep their attention to the vocal tones, so they become more sensitive to vocal tones than their European American counterparts. In terms of response time, for 8-9-year-old Chinese children, the same cultural differences in interferences effect of vocal tones were also significant. However, for the youngest group of children, the interference scores in response time had large variance, which showed large individual differences that may have made the group means non-representative.

Finally, we observed cultural differences between European American and Chinese children, but the two groups of Chinese children performed very similarly in the two tasks. The large cities in China are considerably Westernized. People's life in urban cities in China is infused with elements from Western cultures, such as foreign movies, social media, songs, blogs, international online shopping, and so on. There is no significant difference in material life in big cities between China and the U.S. However, the results of this project suggest that the difference in cultural values remains. Chinese culture may continue to be a high-context culture despite of the influence by Western cultures. People in modern China still put great emphasis on contextual information such as vocal tones, and transmit such cultural values to children through socialization process. This dissertation suggests that cultural differences in values are ingrained and may not be modified easily or quickly. On the other hand, Chinese children in urban and rural samples performed very similarly on the tasks. It is possible that the cultural values are

robust and held by Chinese people regardless of their regions or socioeconomic status. Nevertheless, the lack of urban-rural difference may be also due to the features of the rural sample in this project. The Chinese rural children were recruited in a village near the border of Beijing and Hebei province. Like many other regions in China that were traditionally considered rural, the village is undergoing rapid urbanization. Some people there still live by farming, but many people have sold their farms and work in different positions. Their income has increased significantly and they have got access to resources which were not available for them a decade ago. In 2016, the average annual income was 22,310 RMB (3233 USD) for people in rural areas in Beijing, and 57,275 RMB (8301 USD) for people in urban Beijing. Further studies in villages which have maintained rural during this wave of urbanization will be helpful for understanding the urban-rural difference in detecting emotional cues.

Future directions

One interesting future question is to study whether there are cultural differences in children's recognition of subtle facial expressions. Given that contextual information is valued in Chinese culture, Chinese children may be socialized to be more ready to detect emotions from subtle facial expressions as well. Previous research on emotion recognition primarily focused on prototype facial expression, which are rare to observe in daily life. It would be more practical and meaningful to study children's recognition of subtle facial expressions. Another interesting future direction is to examine the

correlation between children's sensitivity to emotional cues, especially negative emotional cues, and their anxiety level. Previous literature has shown that attention bias toward negative emotional cues from facial expression is positively related to children's anxiety (LoBue & Perez-Edgar, 2014). However, no studies have yet examined whether attention bias to negative emotional tones is a risk factor of anxiety, and whether culture plays a role in such correlations.

In conclusion, this dissertation provides a first demonstration of the cultural similarities and differences in children's sensitivity to emotional cues from word meanings and vocal tones. It sheds light on our knowledge about cultural variations in children's understanding of emotions, and helps researchers delineate a more complete picture of children's emotion knowledge. In addition, this study is useful from a practical perspective. The awareness of cultural differences in emotion processing can facilitate inter-cultural communication. This study is particular important for advising families with multicultural backgrounds to socialize their children with one cultural style, meanwhile to prepare children for socio-emotional competence in other cultural contexts as well. In addition, the results about cultural differences in children's emotion communication provide information to service providers, teachers, medical professionals, and other people who work with children, so that they are able to recognize cultural variations in children's emotion communication and better help children from various cultural backgrounds.

REFERENCES

Adachi, M., Trehub, S. E., & Abe, J.-I. (2004). Perceiving emotion in children's songs across age and culture. *Japanese Psychological Research, 46*(4), 322-336. doi: 10.1111/j.1468-5584.2004.00264.x

Adolphs, R. (2002). Neural systems for recognizing emotion. *Current Opinion in Neurobiology, 12*(2), 169-177. doi: 10.1016/S0959-4388(02)00301-X

Ambady, N., Koo, J., Lee, F., & Rosenthal, R. (1996). More than words: Linguistic and nonlinguistic politeness in two cultures. *Journal of Personality and Social Psychology, 70*(5), 996-1011. doi: 10.1037/0022-3514.70.5.996

Awasthi, A., & Mandal, M. (2015). Facial Expressions of Emotions: Research Perspectives. In M. K. Mandal & A. Awasthi (Eds.), *Understanding Facial Expressions in Communication* (pp. 1-18): Springer India.

Banse, R., & Scherer, K. R. (1996). Acoustic profiles in vocal emotion expression. *Journal of Personality and Social Psychology, 70*(3), 614-636. doi: 10.1037/0022-3514.70.3.614

Buck, R., & Miller, M. (2014). Beyond Facial Expression: Spatial Distance as a Factor in the Communication of Discrete Emotions. *The Social Psychology of Nonverbal Communication*, 173.

Castellano, G., Villalba, S., & Camurri, A. (2007). Recognising Human Emotions from Body Movement and Gesture Dynamics. In A. R. Paiva, R. Prada & R. Picard

(Eds.), *Affective Computing and Intelligent Interaction* (Vol. 4738, pp. 71-82): Springer Berlin Heidelberg.

Campbell, J. R., & Mandel, F. (1990). Connecting math achievement to parental influences. Contemporary Educational Psychology, 15(1), 64-74. doi: 10.1016/0361-476X(90)90006-MCamras, L. A., Oster, H., Campos, J., Campos, R., Ujiie, T., Miyake, K., . . . Meng, Z. (1998). Production of emotional facial expressions in European American, Japanese, and Chinese infants. *Developmental Psychology, 34*(4), 616-628. doi: 10.1037/0012-1649.34.4.616

Chan, S. M., Bowes, J., & Wyver, S. (2009). Chinese parenting in Hong Kong: links among goals, beliefs and styles. *Early Child Development and Care, 179*(7), 849-862. doi: 10.1080/03004430701536525

Chao, R. K. (1995). Chinese and European American Cultural Models of the Self Reflected in Mothers' Childrearing Beliefs. *Ethos, 23*(3), 328-354. doi: 10.1525/eth.1995.23.3.02a00030

Chao, R. K. (1996). Chinese and European American mothers' beliefs about the role of parenting in children's school success. *Journal of Cross-Cultural Psychology, 27*(4), 403-423. doi: 10.1177/0022022196274002

Chen, X. (2000). Growing up in a collectivist culture: Socialization and socioemotional development in Chinese children. In A. L. C. U. P. Gielen (Ed.), *International*

perspectives on human development (pp. 331-353). Lengerich, Germany: Pabst Science Publishers.

Chua, H. F., Boland, J. E., & Nisbett, R. E. (2005). Cultural variation in eye movements during scene perception. *Proceedings of the National Academy of Sciences of the United States of America, 102*(35), 12629-12633. doi: 10.1073/pnas.0506162102

Cole, P. M., & Tan, P. Z. (2007). Emotion socialization from a cultural perspective. *Handbook of socialization: Theory and research*, 516-542.

Denham, S. A. (1986). Social Cognition, Prosocial Behavior, and Emotion in Preschoolers: Contextual Validation. *Child Development, 57*(1), 194-201. doi: 10.2307/1130651

Denham, S. A., Bassett, H. H., & Wyatt, T. (2007). The socialization of emotional competence. In J. E. Grusec (Ed.), & P. D. Hastings (Ed.), Handbook of socialization: Theory and research, (pp. 614-637). New York: Guilford Press.

Denham, S. A., & Couchoud, E. A. (1990). Young preschoolers' understanding of emotions. *Child Study Journal, 20*(3), 171.

Denham, S. A., Renwick-DeBardi, S., & Hewes, S. (1994). Emotional Communication Between Mothers and Preschoolers: Relations With Emotional Competence. *Merrill-Palmer Quarterly, 40*(4), 488-508. doi: 10.2307/23087919

Denham, S. A., Zoller, D., & Couchoud, E. A. (1994). Socialization of preschoolers' emotion understanding. *Developmental Psychology, 30*(6), 928-936. doi: 10.1037/0012-1649.30.6.928

De Sonneville, L. M. J., Verschoor, C. A., Njiokiktjien, C., Op het Veld, V., Toorenaar, N., & Vranken, M. (2002). Facial identity and facial emotions: speed, accuracy, and processing strategies in children and adults. *Journal of Clinical and Experimental Neuropsychology, 24*(2), 200-213. doi: 10.1076/jcen.24.2.200.989

Doan, S. N., & Wang, Q. (2010). Maternal discussions of mental states and behaviors: relations to emotion situation knowledge in European American and immigrant Chinese children. *Child Development, 81*(5), 1490-1503. doi: 10.1111/j.1467-8624.2010.01487.x

Dsilva, M. U., & Whyte, L. O. (1998). Cultural differences in conflict styles: Vietnamese refugees andestablished residents. *Howard Journal of Communications, 9*(1), 57-68. doi: 10.1080/106461798247113

Dunn, J., Brown, J., & Beardsall, L. (1991). Family talk about feeling states and children's later understanding of others' emotions. *Developmental Psychology, 27*(3), 448-455. doi: 10.1037/0012-1649.27.3.448

Dunn, J., Slomkowski, C., Donelan, N., & Herrera, C. (1995). Conflict, Understanding, and Relationships: Developments and Differences in the Preschool Years. *Early Education and Development, 6*(4), 303-316. doi: 10.1207/s15566935eed0604_2

Eisenberg, N., Cumberland, A., & Spinrad, T. L. (1998). Parental socialization of emotion. *Psychological Inquiry, 9*(4), 241-273. doi: 10.1207/s15327965pli0904_1

Ekman, P. (1972). *Universal and cultural differences in facial expression of emotion.* Paper presented at the Nebraska symposium on motivation.

Ekman, P. (1994). Strong evidence for universals in facial expressions: A reply to Russell's mistaken critique. *Psychological Bulletin, 115*(2), 268-287. doi: 10.1037/0033-2909.115.2.268

Ekman, P. (2007). *Emotions revealed: Recognizing faces and feelings to improve communication and emotional life*: Macmillan.

Ekman, P., Friesen, W. V., O'Sullivan, M., Chan, A., Diacoyanni-Tarlatzis, I., Heider, K., . . . Tzavaras, A. (1987). Universals and cultural differences in the judgments of facial expressions of emotion. *Journal of Personality and Social Psychology, 53*(4), 712-717. doi: 10.1037/0022-3514.53.4.712

Elfenbein, H. A., & Ambady, N. (2002). On the universality and cultural specificity of emotion recognition: A meta-analysis. *Psychological Bulletin, 128*(2), 203-235. doi: 10.1037/0033-2909.128.2.203

Fivush, R. (1998). Gendered narratives: Elaboration, structure, and emotion in parent-child reminiscing across the preschool years. In C. P. Thompson, D. J. Herrmann, D. Bruce, J. D. Read, D. G. Payne & M. P. Toglia (Eds.), *Autobiographical*

memory: Theoretical and applied perspectives (pp. 79-104). Mahwah, NJ: Lawrence Erlbaum Associates Publishers.

Fivush, R., Brotman, M., Buckner, J., & Goodman, S. (2000). Gender differences in parent–child emotion narratives. *Sex Roles, 42*(3-4), 233-253. doi: 10.1023/a:1007091207068

Friedlmeier, W., Corapci, F., & Cole, P. M. (2011). Emotion Socialization in Cross-Cultural Perspective. *Social and Personality Psychology Compass, 5*(7), 410-427. doi: 10.1111/j.1751-9004.2011.00362.x

Gao, X., & Maurer, D. (2010). A happy story: Developmental changes in children's sensitivity to facial expressions of varying intensities. *Journal of Experimental Child Psychology, 107*(2), 67-86. doi: 10.1016/j.jecp.2010.05.003

Garner, P. W., Jones, D. C., Gaddy, G., & Rennie, K. M. (1997). Low-income mothers' conversations about emotions and their children's emotional competence. *Social Development, 6*(1), 37-52. doi: 10.1111/j.1467-9507.1997.tb00093.x

Gnepp, J. (1989). Children's use of personal information to understand other people's feelings. *Children's understanding of emotion*, 151-177.

Gosselin, P., & Larocque, C. (2000). Facial morphology and children's categorization of facial expressions of emotions: A comparison between Asian and Caucasian faces. *The Journal of Genetic Psychology, 161*(3), 346-358.

Gottman, J. M., Katz, L. F., & Hooven, C. (1997). *Meta-emotion: How families communicate emotionally*: Psychology Press.

Gross, D., & Harris, P. L. (1988). False Beliefs About Emotion: Children's Understanding of Misleading Emotional Displays. *International Journal of Behavioral Development, 11*(4), 475-488. doi: 10.1177/016502548801100406

Halberstadt, A. G., Fox, N. A., & Jones, N. A. (1993). Do expressive mothers have expressive children? The role of socialization in children's affect expression. *Social Development, 2*(1), 48-65. doi: 10.1111/j.1467-9507.1993.tb00004.x

Hall, E. T. (1976). *Beyond culture*. New York: Anchor Books.

Ishii, K., Reyes, J. A., & Kitayama, S. (2003). Spontaneous attention to word content versus emotional tone: Differences among three cultures. *Psychological science, 14*(1), 39-46. doi: 10.1111/1467-9280.01416. Izard, C. E. (1994). Innate and universal facial expressions: Evidence from developmental and cross-cultural research. *Psychological Bulletin, 115*(2), 288-299. doi: 10.1037/0033-2909.115.2.288

Kashima, E. S., & Kashima, Y. (1998). Culture and Language: The Case of Cultural Dimensionsand Personal Pronoun Use. *Journal of Cross-Cultural Psychology, 29*(3), 461-486. doi: 10.1177/0022022198293005

Kessous, L., Castellano, G., & Caridakis, G. (2010). Multimodal emotion recognition in speech-based interaction using facial expression, body gesture and acoustic

analysis. *Journal on Multimodal User Interfaces, 3*(1-2), 33-48. doi: 10.1007/s12193-009-0025-5

Kim, D., Pan, Y., & Park, H. S. (1998). High- Versus Low-Context Culture: A Comparison of Chinese, Korean, and American Cultures. *Psychology & Marketing (1986-1998), 15*(6), 507.

Kipp, M., & Martin, J. C. (2009, 10-12 Sept. 2009). *Gesture and emotion: Can basic gestural form features discriminate emotions?* Paper presented at the Affective Computing and Intelligent Interaction and Workshops, 2009. ACII 2009. 3rd International Conference on Affective Computing and Intelligent Interaction. IEEE Press.

Kitayama, S., Duffy, S., Kawamura, T., & Larsen, J. T. (2003). Perceiving an object and its context in different cultures: A cultural look at new look. *Psychological Science, 14*(3), 201-206. doi: 10.1111/1467-9280.02432

Kitayama, S., & Ishii, K. (2002). Word and voice: Spontaneous attention to emotional utterances in two languages. *Cognition and Emotion, 16*(1), 29-59. doi: 10.1080/0269993943000121

Leets, L. (2003). Disentangling Perceptions of Subtle Racist Speech A Cultural Perspective. *Journal of Language and Social Psychology, 22*(2), 145-168. doi: 10.1177/0261927X03022002001

Lagattuta, K. H., & Thompson, R. A. (2007). The development of self-conscious emotions: Cognitive processes and social influences. In J. L. Tracy, R. W. Robins & J. P. Tangney (Eds.), *The self-conscious emotions: Theory and research* (pp. 91-113). New York: Guilford Press.

Lewis, M., Ramsay, D. S., & Kawakami, K. (1993). Differences between Japanese Infants and Caucasian American Infants in Behavioral and Cortisol Response to Inoculation. *Child Development, 64*(6), 1722-1731. doi: 10.1111/j.1467-8624.1993.tb04209.x

LoBue, V., & Pérez‐Edgar, K. (2014). Sensitivity to social and non‐social threats in temperamentally shy children at‐risk for anxiety. *Developmental science, 17*(2), 239-247. doi: 10.1111/desc.12110

Lu, Z., Daneman, M., & Reingold, E. M. (2008). Cultural differences in cognitive processing style: Evidence from eye movements during scene processing.

Markham, R., & Wang, L. (1996). Recognition of emotion by Chinese and Australian children. *Journal of Cross-Cultural Psychology, 27*(5), 616-643. doi: 10.1177/0022022196275008

Markus, H. R., & Kitayama, S. (1991). Culture and the self: Implications for cognition, emotion, and motivation. *Psychological Review, 98*(2), 224-253. doi: 10.1037/0033-295x.98.2.224

Markus, H. R., & Kitayama, S. (2001). The cultural construction of self and emotion: Implications for social behavior. *Emotions in social psychology: Essential reading*, 119-137.

Masuda, T., Ellsworth, P. C., Mesquita, B., Leu, J., Tanida, S., & Van de Veerdonk, E. (2008). Placing the face in context: Cultural differences in the perception of facial emotion. *Journal of Personality and Social Psychology, 94*(3), 365-381. doi: 10.1037/0022-3514.94.3.365

Masuda, T., & Nisbett, R. E. (2001). Attending holistically versus analytically: Comparing the context sensitivity of Japanese and Americans. *Journal of Personality and Social Psychology, 81*(5), 922-934. doi: 10.1037/0022-3514.81.5.922

Matsumoto, D. (1991). Cultural influences on facial expressions of emotion. *Southern Communication Journal, 56*(2), 128-137. doi: 10.1080/10417949109372824

Matsumoto, D., & Willingham, B. (2009). Spontaneous facial expressions of emotion of congenitally and noncongenitally blind individuals. *Journal of Personality and Social Psychology, 96*(1), 1-10. doi: 10.1037/a0014037

Nisbett, R. E., Peng, K., Choi, I., & Norenzayan, A. (2001). Culture and systems of thought: Holistic versus analytic cognition. *Psychological Review, 108*(2), 291-310. doi: 10.1037/0033-295x.108.2.291

Pell, M. D., & Kotz, S. A. (2011). On the Time Course of Vocal Emotion Recognition. *PLoS ONE, 6*(11), e27256. doi: 10.1371/journal.pone.0027256

Porter, C. L., Hart, C. H., Yang, C., Robinson, C. C., Olsen, S. F., Zeng, Q., . . . Jin, S. (2005). A comparative study of child temperament and parenting in Beijing, China and the western United States. *International Journal of Behavioral Development, 29*(6), 541-551. doi: 10.1177/01650250500147402

Ramsey, S. J. (1998). Interactions between North Americans and Japanese: Considerations of communication style. *Basic concepts of intercultural communication*, 111-130.

Rigoulot, S., Wassiliwizky, E., & Pell, M. D. (2013). Feeling backwards? How temporal order in speech affects the time course of vocal emotion recognition. *Frontiers in Psychology, 4*, 367. doi: 10.3389/fpsyg.2013.00367

Saarni, C., Campos, J. J., Camras, L. A., & Witherington, D. (2006). Emotional development: Action, communication, and understanding (N. Eisenberg, Trans.) *Social, emotional, and personality development* (6th ed., Vol. 3, pp. 226-299). Hoboken, NJ: John Wiley & Sons. Inc.

Salleh, L. M. (2005). *High/low context communication: The Malaysian Malay style.* Paper presented at the Proceedings of the 2005 Association for Business Communication Annual Convention.

Scherer, K. R. (1999). Universality of emotional expression. *Encyclopedia of human emotions, 2*, 669-674.

Scherer, K. R. (2003). Vocal communication of emotion: A review of research paradigms. *Speech Communication, 40*(1–2), 227-256. doi: 10.1016/S0167-6393(02)00084-5

Scherer, K. R., Banse, R., & Wallbott, H. G. (2001). Emotion Inferences from Vocal Expression Correlate Across Languages and Cultures. *Journal of Cross-Cultural Psychology, 32*(1), 76-92. doi: 10.1177/0022022101032001009

Senzaki, S. (2013). *The development and transmission of culturally unique attentional styles in Canada and Japan: A demonstration of children's cultural learning and parents' scaffolding behaviors.* University of Alberta.

Suveg, C., Zeman, J., Flannery-Schroeder, E., & Cassano, M. (2005). Emotion Socialization in Families of Children With an Anxiety Disorder. *Journal of Abnormal Child Psychology, 33*(2), 145-155. doi: 10.1007/s10802-005-1823-1

Tao, A., Zhou, Q., & Wang, Y. (2010). Parental reactions to children's negative emotions: Prospective relations to Chinese children's psychological adjustment. *Journal of Family Psychology, 24*(2), 135-144. doi: 10.1037/a0018974

Thompson, R. A. (1989). Causal attributions and children's emotional understanding. In C. Saarni & P. Harris (Eds.), *Children's understanding of emotion* (pp. 117-150). New York: Cambridge University Press.

Trommsdorff, G., & Friedlmeier, W. (2009). Preschool girls' distress and mothers' sensitivity in Japan and Germany. *European Journal of Developmental Psychology, 7*(3), 350-370. doi: 10.1080/17405620802252742

Valiente, C., Eisenberg, N., Fabes, R. A., Shepard, S. A., Cumberland, A., & Losoya, S. H. (2004). Prediction of Children's Empathy-Related Responding From Their Effortful Control and Parents' Expressivity. *Developmental Psychology, 40*(6), 911-926. doi: 10.1037/0012-1649.40.6.911

Wang, Q. (2001). “Did you have fun?”: American and Chinese mother–child conversations about shared emotional experiences. *Cognitive Development, 16*(2), 693-715. doi: 10.1016/S0885-2014(01)00055-7

Wang, Q. (2003). Emotion situation knowledge in American and Chinese preschool children and adults. *Cognition and Emotion, 17*(5), 725-746. doi: 10.1080/02699930302285

Wang, Q. (2008). Emotion knowledge and autobiographical memory across the preschool years: A cross-cultural longitudinal investigation. *Cognition, 108*(1), 117-135. doi: 10.1016/j.cognition.2008.02.002

Wang, Q. (2013). Chinese socialization and emotion talk between mothers and children in native and immigrant Chinese families. *Asian American Journal of Psychology, 4*(3), 185-192. doi: 10.1037/a0030868

Wang, Q., & Fivush, R. (2005). Mother–Child Conversations of Emotionally Salient Events: Exploring the Functions of Emotional Reminiscing in European-American and Chinese Families. *Social Development, 14*(3), 473-495. doi: 10.1111/j.1467-9507.2005.00312.x

Wang, Q., Hutt, R., Kulkofsky, S., McDermott, M., & Wei, R. (2006). Emotion situation knowledge and autobiographical memory in Chinese, immigrant Chinese, and European American 3-year-olds. *Journal of Cognition and Development, 7*(1), 95-118. doi: 10.1207/s15327647jcd0701_5

Wang, Q., Leichtman, M. D., & Davies, K. I. (2000). Sharing memories and telling stories: American and Chinese mothers and their 3-year-olds. *Memory, 8*(3), 159-177. doi: 10.1080/096582100387588

Widen, S. C., & Russell, J. A. (2010). Children's scripts for social emotions: Causes and consequences are more central than are facial expressions. *British Journal of Developmental Psychology, 28*(3), 565-581. doi: 10.1348/026151009X457550d

Yang, Y. & Wang, Q. (2016). The relation of emotion knowledge to coping in European American and Chinese immigrant children. Journal of Child and Family Studies, 25(2), 452-463. Doi: 10.1007/s10826-015-0224-5.Louie, J. Y., Oh, B. J., & Lau, A. S. (2013). Cultural differences in the links between parental control and children's emotional expressivity. *Cultural Diversity and Ethnic Minority Psychology, 19*(4), 424-434. doi: 10.1037/a0032820

Yanushevskaya, I., Gobl, C., & Ní Chasaide, A. (2013). Voice quality in affect cueing: does loudness matter? *Frontiers in Psychology, 4*, 335. doi: 10.3389/fpsyg.2013.00335

Zahn-Waxler, C., Friedman, R. J., Cole, P. M., Mizuta, I., & Hiruma, N. (1996). Japanese and United States preschool children's responses to conflict and distress. *Child Development, 67*(5), 2462-2477. doi: 10.1111/j.1467-8624.1996.tb01868.x

APPENDICE

Appendix 1: Instructions on the meaning task and tone task (English).

"We are going to play a fun computer game. Do you like computer games? Well, I hope you like this game. Before we play I want you to put on these headphones. But, if they are too loud or are uncomfortable, please let me know so I can fix it." (if child doesn't want to put on headphones then ask them whether earbuds are better for them).

Meaning game:

"The first game is MEANING GAME. In this game, you are going to hear some words. Some words have good meanings, like happy or nice, and some words have bad meanings, like mean or mad. When you hear a good word, I want you to press the smiley button. And when you hear a bad word I want you to press the frowny button. After hearing a word, press the button as fast as you can. In this game, you should ignore the

vocal tone of the word you hear, that is how a word is said, and focus on the word meaning. Ready? Let's give it a try!"

Make sure the child has one finger by each button and is ready to begin before starting.

Practice trials (8):

Children get feedback on each trial. If correct, children will see a green √ on the screen and hear a positive sound "Di;" if incorrect, children will see a red question mark on the screen and hear a negative sound "Zzzz." The experimenter reminds the participant: xxx has a good/bad meaning, so you should press the smiley/frowny button.

After all practice trials:

"Very good! Now you are ready for the real game! Remember, this is the meaning game. Press the smiley button when you hear a good word, and press the frowny button when you hear a bad word. Ignore the sound and focus on the meaning. Press as fast as you can"

Test trials (24):

Children won't get feedback. Response latency will be recorded after the word starts. The next beep starts at 1500ms after the response.

Tone game:

"Now, you will play another game. It's the TONE GAME. Do you know what tone means? For example, if I say "xxx" (call the participant's name with a happy tone), that

sounds happy, right? If I say “xxx” (call the participant’s name with a sad tone), that sounds sad, right? In this game, you are going to hear some words again. Some words sound happy, and some words sound sad or mad. When you hear a happy voice or happy sound I want you to press the smiley button. And when you hear a sad or mad voice, I want you to press the frowny button. After you hear a word, press the button as fast as you can. In this game, you need to ignore the meaning of the word you hear, that is what a word actually means, and focus on the vocal tone. Ready? Let’s try!”

Practice trials (8):

Children get feedback on each trial. If correct, children will see a green √ on the screen and hear a positive sound “Di;” if incorrect, children will see a red question mark on the screen and hear a negative sound “Zzzz.” The experimenter reminds the participant: xxx sounds happy/sad/mad, so you should press the smiley/frowny button.

After all practice trials:

“Very good! Now you are ready for the real game! Remember, this is the TONE game. Press the smiley button when you hear happy voice, and press the frowny button when you hear sad or mad voice. Ignore the meaning of the word and focus on the sound. Press as fast as you can”

Test trials (24):

Children won’t get feedback. Response latency will be recorded after the word starts. The next beep starts at 1500ms after the response.

Say “good job” to the child after each block (practice and test), give a sticker after each test block.

The order of two games is counter balanced. The order of practice trials is fixed, but the order of test trials is randomized.

Appendix 2: Instructions on the meaning task and tone task (Chinese).

“我们来玩一个电脑游戏。你喜欢玩游戏吗？好，希望你喜欢。开始游戏之前，我们先把耳机带上吧。如果声音太大或者带着不舒服，就告诉我，我会帮你弄一下。”（如果被试不想带耳机，问问他换另一种耳机会不会好点）

意思游戏：

“我们今天玩的第一个游戏是“意思游戏”。在这个游戏里，你会听到一些词，有的词有好的意思，比如高兴、好人，有的词有不好的意思，比如生气、坏人。当你听到一个好词，你要按带有笑脸的这个按钮；当你听到一个不好的词，你要按带有哭脸的这个按钮。听完一个词之后越快按越好。在这个游戏中，你不要去管这些词的语气，不去管词是怎么说的，只根据意思判断。准备好了吗？我们来试一下。”

开始之前，确保被试在每个按键上都放了一个手指。

练习：（12）

被试在每个试次都得到反馈。如果正确，屏幕上会显示绿色对勾，声音反馈为“嘀” 如果，回答错误，屏幕上会显示红色问号，声音反馈为“当！” 实验者告

诉被试：”记得听到好词按笑脸，听到坏词按哭脸。Xx 这个词有好/不好的意思对吧，所以应该按笑/哭脸”

练习结束后，“非常棒！我们来正式开始游戏吧！记住这是意思游戏，听到好的词按笑脸，不好的词按哭脸，不去管声音是什么样的，听完一个词越快按越好”

测试：（36）

无反馈。

测试结束后：“非常棒，奖励你一个贴画。”

语气游戏：

“现在，我们要玩另一个游戏，这是个“语气游戏”。你知道什么是语气吗？比如我说，“xxx”(用高兴的语气叫被试的名字)，这样是不是听起来很高兴？这是高兴的语气，高兴的声音。如果我说，”xxx”（用不高兴的语气叫被试的名字），这样是不是听起来不高兴？这是不高兴的语气，不高兴的声音。在这个游戏里，你还会听到一些词，有些词听起来很高兴，有些词听起来不高兴。我想让你听到高兴的声音按笑脸，听到不高兴的声音按哭脸。听完一个词之后越快按越好。在这个游戏里，不要去管这个词意思是什么。只根据声音来判断。准备好了吗？我们来试一下。”

练习：（12）

被试在每个试次都得到反馈。如果正确，屏幕上会显示绿色对勾，声音反馈为嘀；如果回答错误，屏幕上会显示红色问号，声音反馈为当！实验者告诉被试：” 记得听到高兴的语气按笑脸，刚才听到的有好/不好的语气对吧，所以应该按笑/哭脸。”

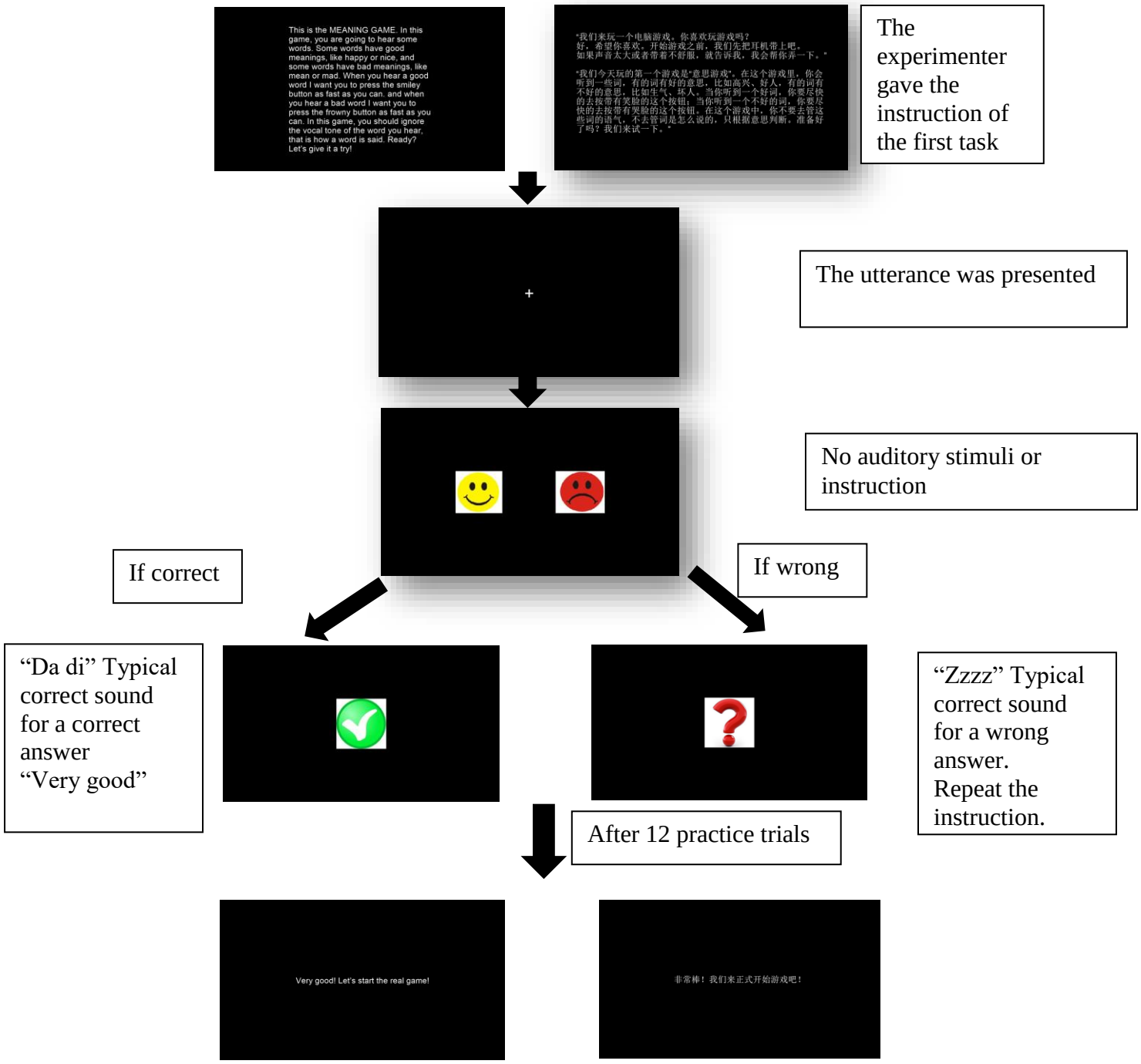
练习结束后，“非常棒！我们来正式开始游戏吧！记得这是语气游戏，高兴的声音按笑脸，不高兴的声音按哭脸，不去管意思是什么，听完一个词之后越快按越好。”

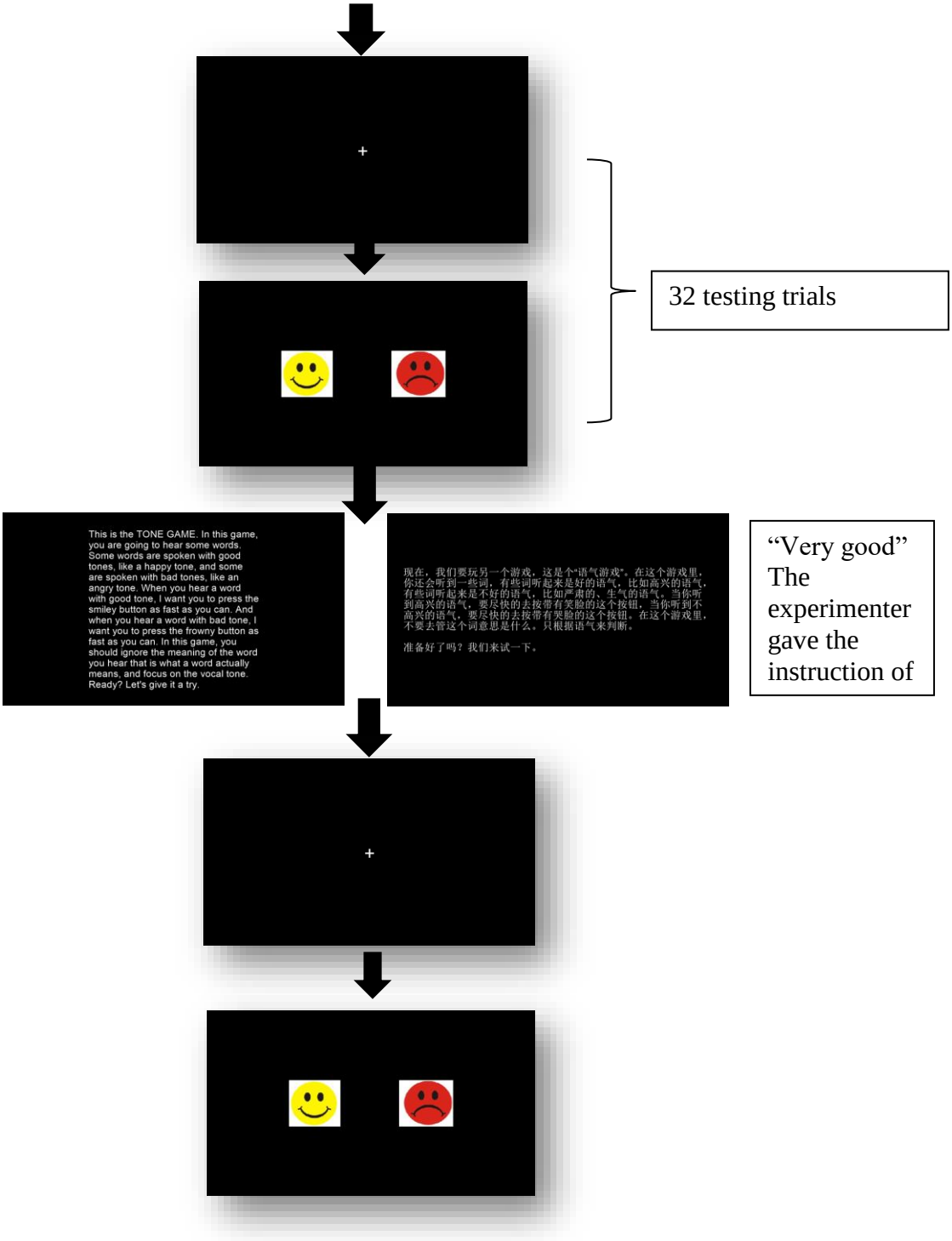
测试：（36）

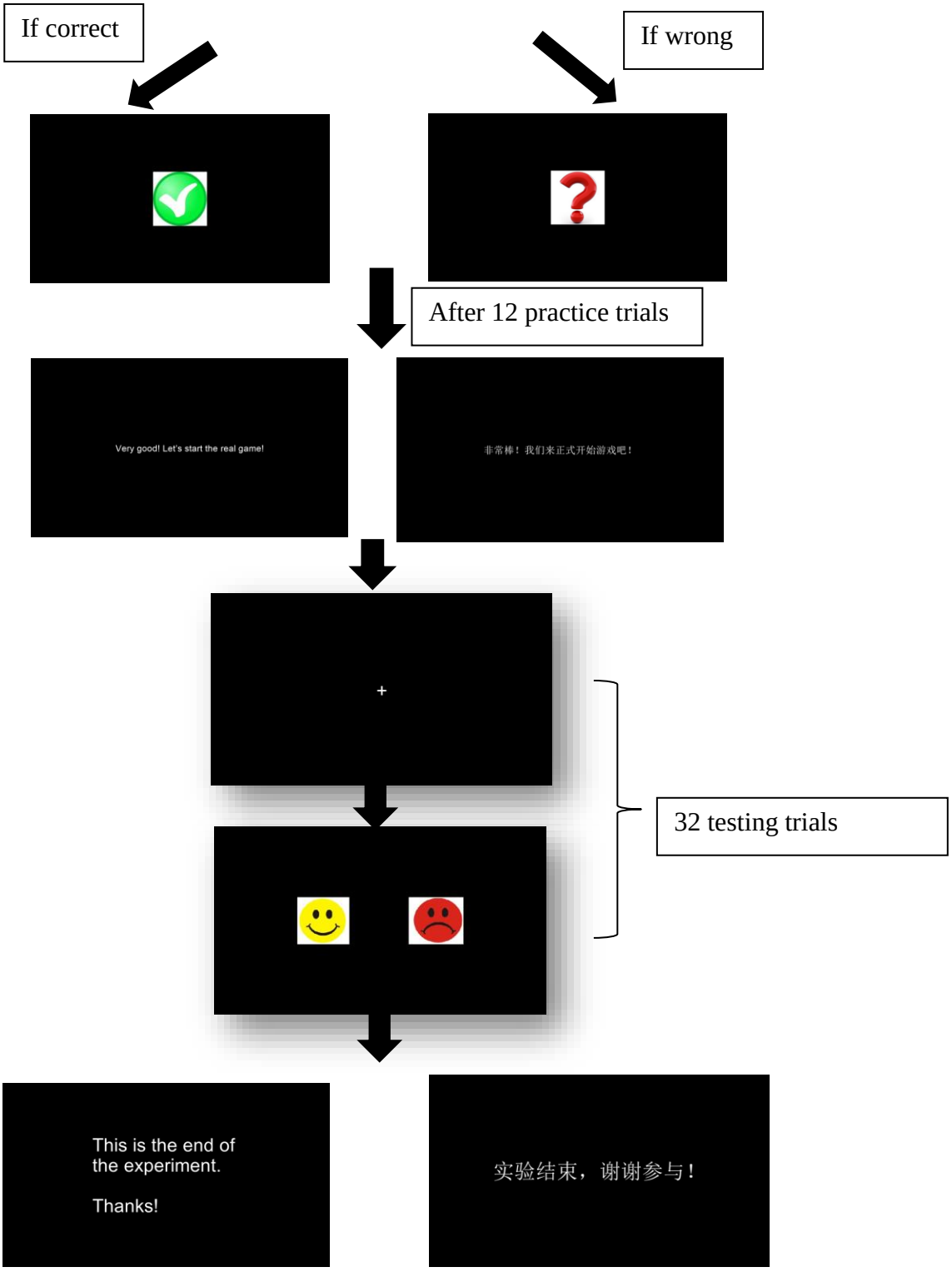
无反馈。

测试结束后：“非常棒，再奖励你一个贴画。”

Appendix 3. Screenshots of the experiment.




+
32 testing trials
This is the TONE GAME. In this game, you are going to hear some words. Some words are spoken with good tones, like a happy tone, and some are spoken with bad tones, like an angry tone. When you hear a word with good tone, I want you to press the smiley button as fast as you can. And when you hear a word with bad tone, I want you to press the frowny button as fast as you can. In this game, you should ignore the meaning of the word you hear that is what a word actually means, and focus on the vocal tone. Ready? Let's give it a try.
现在，我们要玩另一个游戏，这是个"语气游戏"。在这个游戏里，你还会听到一些词，有些词听起来是好的语气，比如高兴的语气，有些词听起来是不好的语气，比如严肃的、生气的语气。当你听到高兴的语气，要尽快的去按带有笑脸的这个按钮，当你听到不高兴的语气，要尽快的去按带有哭脸的这个按钮。在这个游戏里，不要去管这个词意思是什么。只根据语气来判断。
准备好了吗？我们来试一下。
"Very good" The experimenter gave the instruction of
+


If correct
If wrong
After 12 practice trials
Very good! Let's start the real game!
非常棒！我们来正式开始游戏吧！
+
32 testing trials
This is the end of
the experiment.
Thanks!
实验结束，谢谢参与！

Appendix 4. Picture of setup.